Davinia Hernández-Leo
Tobias Ley
Ralf Klamma
Andreas Harrer (Eds.)

LNCS 8095

Scaling up Learning for Sustained Impact

8th European Conference on Technology Enhanced Learning, EC-TEL 2013
Paphos, Cyprus, September 2013, Proceedings

Springer

Lecture Notes in Computer Science 8095

Davinia Hernández-Leo Tobias Ley
Ralf Klamma Andreas Harrer (Eds.)

Scaling up Learning for Sustained Impact

8th European Conference
on Technology Enhanced Learning, EC-TEL 2013
Paphos, Cyprus, September 17-21, 2013
Proceedings

Springer

Volume Editors

Davinia Hernández-Leo
Universitat Pompeu Fabra, Barcelona, Spain
E-mail: davinia.hernandez@upf.edu

Tobias Ley
Tallinn University, Estonia
E-mail: tley@tlu.ee

Ralf Klamma
RWTH Aachen University, Germany
E-mail: klamma@dbis.rwth-aachen.de

Andreas Harrer
Catholic University Eichstätt-Ingolstadt, Germany
E-mail: andreas.harrer@ku-eichstaett.de

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-40813-7 e-ISBN 978-3-642-40814-4
DOI 10.1007/978-3-642-40814-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013947074

CR Subject Classification (1998): H.4, H.3, I.2, C.2, H.5, J.1

LNCS Sublibrary: SL 2 – Programming and Software Engineering

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

These proceedings collect the papers presented at the Eighth European Conference on Technology Enhanced Learning (EC-TEL) held in Paphos (Cyprus) during September 17–21, 2013. In this edition, the interdisciplinary TEL community in Europe and Worldwide - has again recognized the conference as a key venue for discussing and disseminating research outcomes in the field. Since 2006, EC-TEL has been a reference scene for relevant state-of-the-art research in TEL; first in Crete (Greece also in 2007), and then in Maastricht (the Netherlands 2008), Nice (France 2009), Barcelona (Spain 2010), Palermo (Italy 2011), and Saarbrücken (Germany 2012).

The theme of this year's conference focused on the need of "Scaling up Learning for Sustained Impact". TEL research outcomes have shown to be able to achieve significant improvements in schools, universities, and at the workplaces. However, these innovations have tended to be unsustainable. Too often, initiatives need a high degree of effort to be sustained, and are consequently endangered when funding stops. At the same time, mobile and social information and communication technologies are fundamentally changing our lives. Nowadays we use them naturally to keep in contact with our friends, to seek information, to buy things and to work. In many domains, these emerging technologies demonstrate that scaling is possible, but few educational institutions or companies have taken these technologies up in a systematic way to include them into their learning strategy. The question of how to scale up learning is not trivial. To be successful, scaling must be obtained on a number of different levels. Some of the research questions that need to be addressed include:

- How can next generation social and mobile technologies be used to support informal learning processes?
- How can TEL be better integrated within our daily work and learning practices?
- What are the pedagogical and technological challenges of massive open learning scenarios and how can TEL research address them?
- How can TEL support complex learning ecosystems involving multiple devices, activities, resources, and agents?
- How can novel architectural solutions and social-semantic technologies help us scale solutions beyond the immediate learning context?

Drawing on the core TEL disciplines of Computer Science, Network and Communication Technologies, Education, Psychology, Cognitive Science, and Social Science, research contributions presented in EC-TEL 2013 have tackled these questions. Topics addressed include Open Educational Resources (OER), Massive Open Online Courses (MOOC), Schools of the Future, Orchestration of Learning Activities, Learning Networks, Teacher Networks, Bring your own Device (BYOD), Social Media, Learning Analytics, Personalization, Mobile Learning, Computer-Supported Collaborative Learning, Game-based and Simulation-based Learning, and Learning Design.

This 2013 edition has been extremely competitive, given the record number of submissions generated. Out of the more than 200 initial abstract submissions, a total of 194 valid paper submissions were received. 136 of them were full papers. All submissions were assigned to three members of the Program Committee (PC) for review. One of the reviewers had the role of leading reviewer and initiated a discussion in case of conflicting reviews. All reviews as well as the discussions were checked and discussed within the team of PC chairs, and additional reviews or metareviews were elicited if necessary. Finally, 31 submissions were selected as full papers (which resulted in an acceptance rate for full papers of 22.8%). Additionally, 18 papers were presented as short papers, 14 as demonstrations and 29 as posters. The dedicated work of all the PC members as well as the additional reviewers must be acknowledged. Only with their help was it possible to deal with the record number of submissions and still keep all deadlines as originally planned.

Keynote presentations completed this competitive scientific program, Sir John Daniel, one of the world's most eminent practitioners and thinkers in open, distance, and technology mediated learning, and Ms Stamenka Uvalic-Trumbic, Former Chief of the Higher Education Section of the United Nations Organization for Education, Science and Culture (UNESCO), gave a talk entitled "Making Sense of MOOCs: The Evolution of Online Learning in Higher Education". An extended abstract of their talk is included in these proceedings. Following the conference tradition, the European Commission was also present in EC-TEL 2013 to introduce the strategic priorities and directions in the field. This year, Liina Munari, senior research project officer at the European Commission, DG Connect, Unit G4 "Youth, Skills and Inclusion", gave a talk with the title "Beyond Horizon: learning and technology in EC policies and programming".

Continuing with the tradition started in EC-TEL 2012, demonstrations had a pronounced role in the conference program. A plenary session was organized as a "TEL Demo sprint contest" in which the demonstrations were shown "in action" giving the audience the possibility to vote for the best demo. Demonstrations and posters were also interactively exhibited during the conference, sparking discussions between different but complementary groups, bringing the community closer to solving the manifold problems we are facing. Besides, representatives from the industry presented and discussed their vision of the field in the Industry Track.

The TEL community also proposed and organized a number of stimulating workshops as part of the conference. 10 workshops and tutorials were organized, ranging from Serious Games to Workplace Learning, addressing some particular domains, such as Health Care, or focusing on cross-cutting issues in TEL, such as motivation or reflection. A Doctoral Consortium was organized concurrently with the workshops. It provided an opportunity for PhD students to discuss their work with experienced TEL researchers. Finally, a soccer tournament completed the program as an informal community-building activity.

Many contributions have made the conference possible: the enthusiastic authors, the supportive PC members, and the diligent Chairs forming the Executive Committee. The conference has also been partially supported in one way or another by the European Association of Technology-Enhanced Learning (EATEL), the University of Cyprus, Online Educa Berlin (sponsor of the EC-TEL2013 Demo Award), Springer, EasyChair, and the IEEE Transactions on Learning Technologies.

EC-TEL has always acknowledged the need for sustaining the impact of TEL research outcomes. This year the potential of "scalability" has been brought into focus as a means of reaching the required sustainability. Hopefully, while reading these proceedings, you will also recognize the relevance of this perspective in TEL and further contribute to "scale up" the herewith-reported research efforts "for their sustained and enhanced impact".

September 2013

Davinia Hernández-Leo
Tobias Ley
Ralf Klamma
George Papadopoulos

Organization

Executive Committee

General Chair

Ralf Klamma	RWTH Aachen University, Germany

Program Chairs

Davinia Hernández-Leo	Universitat Pompeu Fabra, Barcelona, Spain
Tobias Ley	Tallinn University, Estonia

Industry Chair

Lucia Pannese	imaginary, Italy

Workshop Chair

Anoush Margaryan	Glasgow Caledonian University, UK

Poster and Demonstration Chair

Andreas Harrer	Catholic University Eichstätt-Ingolstadt, Germany

Doctoral Consortium Chairs

Katherine Maillet	Tèlècom SudParis, France
Tomaž Klobučar	Jozef Stefan Institute, Slovenia

Dissemination Chair

Michael Derntl	RWTH Aachen University, Germany

Local Organization Chair

George Angelos Papadopoulos	University of Cyprus

Program Committee

Achilleas Achilleos
Mario Allegra
Heidrun Allert
Charoula Angeli
Inmaculada Arnedillo-Sanchez
Miguel Rodriguez Artacho
Merja Bauters
Francesco Bellotti
Adriana Berlanga
Katrin Borcea-Pfitzmann
Francis Brouns
Peter Brusilovsky
Daniel Burgos
Manuel Caeiro

Lorenzo Cantoni
Linda Castañeda
Manuel Castro
Mohamed Amine Chatti
Giuseppe Chiazzese
Agnieszka Chrzaszcz
John Cook
Audrey Cooke
Raquel M. Crespo
Ulrike Cress
Alexandra I Cristea
Paul de Bra
Stavros Demetriadis
Christian Depover
Michael Derntl
Philippe Dessus
Darina Dicheva
Stefan Dietze
Yannis Dimitriadis
Vania Dimitrova
Juan Manuel Dodero
Peter Dolog
Hendrik Drachsler
Jon Dron
Benedict Du Boulay
Erik Duval
Martin Ebner
Alfred Essa
Baltasar Fernandez-Manjon
Carmen Fernández-Panadero
Katharina Freitag
Muriel Garreta
Dragan Gasevic
Sébastien George
Denis Gillet
Fabrizio Giorgini
Carlo Giovannella
Christian Glahn
Sabine Graf
Monique Grandbastien
Andrina Granic
David Griffiths
Begona Gros
Christian Guetl
Joerg Haake
Andreas Harrer
Davinia Hernández-Leo
Knut Hinkelmann
Ulrich Hoppe
Raija Hämäläinen
Christoph Igel
Seiji Isotani
Perez Juan Asensio
Sanna Järvelä
Marco Kalz
Nuri Kara
Nikos Karacapilidis
Michael Kickmeier-Rust
Barbara Kieslinger
Dr Kinshuk
Ralf Klamma
Joris Klerkx
Tomaz Klobucar
Carlos Delgado Kloos
Milos Kravcik
Barbara Kump
Mart Laanpere
Lydia Lau
Elise Lavoue
Effie Law
Tobias Ley
Stefanie Lindstaedt
Andreas Lingnau
Martin Llamas-Nistal
Rose Luckin
George Magoulas
Katherine Maillet
Nikos Manouselis
Anoush Margaryan
Martin Memmel
Christos Mettouris
Riichiro Mizoguchi
Paola Monachesi
Yishay Mor
Pablo Moreno-Ger
Pedro J. Muñoz Merino
Mario Muñoz-Organero
Kati Mäkitalo-Siegl
Rob Nadolski
Ambjorn Naeve

Roger Nkambou
Viktoria Pammer
Lucia Pannese
George Papadopoulos
Abelardo Pardo
Kai Pata
Jermann Patrick
Jan Pawlowski
Niels Pinkwart
Eric Ras
Andrew Ravenscroft
Cristoph Rensing
Luis Anido Rifon
Verónica Rivera-Pelayo
Demetrios Sampson
Andreas Schmidt
J.I. Schoonenboom
Ulrik Schroeder
Mike Sharples
Bernd Simon
Peter Sloep
Mohammad Smadi
Sergey Sosnovsky
Marcus Specht
Daniel Spikol
Slavi Stoyanov
Pierre Tchounikine
Stefaan Ternier
Vladimir Tomberg
Christoph Trattner
Stefan Trausan-Matu
Aimilia Tzanavari
Martin Valcke
Christine Vanoirbeek
Katrien Verbert
Barbara Wasson
Armin Weinberger
Wim Westera
Fridolin Wild
Martin Wolpers
Volker Zimmermann

Additional Reviewers

Abel, Marie-Helene
Ali, Liaqat
Arrigo, Marco
Bernard, Jason
Bientzle, Martina
Bourda, Yolaine
Buis, Stan
Dascalu, Mihai
De Ascaniis, Silvia
de Vreeze, Jort
Draaijer, Silvester
El-Bishouty, Moushir M.
Florian, Beatriz
Gentile, Manuel
Giemza, Adam
Halatchliyski, Iassen
Hansen, Cecilie
Imran, Hazra
Kapitsaki, Georgia
Kovanovic, Vitomir
Latishev, Alexey
Lau, Lydia
Lee, Danielle H.
Leony, Derick
Lezcano, Leonardo
Melero-Gallardo, Javier
Nakic, Jelena
Oeberst, Aileen
Papadopoulos, Pantelis
Petrushyna, Zinayida
Prilla, Michael
Pérez-Sanagustín, Mar
Robles, Luz Stella
Santos, Patricia
Seitlinger, Paul
Siadaty, Melody
Tammets, Kairit
Tardini, Stefano
Tatar, Deborah
Väljataga, Terje
Wake, Jo
Ziebarth, Sabrina

Table of Contents

Invited Paper

Full Papers

Short Papers

Demo Papers

Poster Papers

Making Sense of MOOCs: The Evolution of Online Learning in Higher Education

Stamenka Uvalić-Trumbić and John Daniel

Senior Advisors, Academic Partnerships

1 Introduction

To make sense of MOOCs and understand the evolution of online learning in higher education we must go back to times long before the Internet was invented. Only in this way can we understand the gradual evolution of open, distance and online learning in higher education. We shall look first at the evolution of methods for teaching at a distance and then at how the choice of the content of learning has increased.

2 Teaching at a Distance

Saint Paul was not a university professor. However, his letters to the young churches around the Mediterranean were an early form of distance learning using letters and donkeys as the media to carry the message.

Priests would read out the letters and provide commentary, perhaps with discussion. Saint Paul's system was open. There were no barriers to attending church and engaging with his thinking – unless you were discouraged by the danger of being thrown to the lions during one of the periodic Roman crackdowns on Christianity.

This became a powerful educational movement. We can argue about the relative importance of Saint Paul to the worldwide spread of Christianity, but without question his early system of correspondence education gave the Church doctrinal consistency and later, with a new technology, doctrinal controversy. That new technology, printing, came over a millennium later. The fastidious hand copying of manuscripts was no long necessary. The written word came directly into the hands of ordinary people.

By giving the written word to individuals, printing introduced another important concept in the evolution of education across space and time: independent study. People could now make up their own minds about what a book meant, which led to the upheavals of the Protestant Reformation and less deference to authority. There is a charming Korean saying that there is no nicer sound than the rustle of turning pages as someone reads a book late at night. The broad impact of printing was a first example of the truth of the famous statement by Marshall McLuhan that 'the medium is the message'.

The next key technological advance made the medium of print much more powerful. This was the development of railway networks in the nineteenth century.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 1–4, 2013.

You could now move print rapidly and reliably over distance. Postal services were transformed. When postal systems allowed documents to be exchanged more readily, education reacted quickly. The Penny Post, the first universal postal service, was introduced in Britain in 1840. Isaac Pitman exploited it immediately to teach shorthand by correspondence. He launched the commercial correspondence education industry, which defined distance education for more than a century.

In the 20th century various new technologies came and stayed: radio, film, television, computing and computer assisted learning. Enthusiasts predicted that each new medium would revolutionise education. In 1940 the motion picture was hailed as the most revolutionary instrument introduced into education since the printing press. In 1962 programmed learning was the first major technological innovation since the invention of printing. Not long afterwards it was the impact of computers. Note that these prophets all took printing as their touchstone, not the previous technological marvel.

Wise practitioners conclude from this story that there is no magic educational medium and doubt that there ever will be. No single technology is revolutionary but a combination can be. By the 1960s, the blending of technologies had begun to create a rich communications environment.

At the foundation ceremony of the UK Open University in 1969 the Chancellor, Lord Crowther, captured this in these words: "The world is caught in a communications revolution, the effects of which will go beyond those of the industrial revolution of two centuries ago. Then the great advance was the invention of machines to multiply the potency of men's muscles. Now the great new advance is the invention of machines to multiply the potency of men's minds. As the steam engine was to the first revolution, so the computer is to the second."

It is hard to overstate the impact of the UK Open University. Established with strong political support it created a new synthesis of the technological, pedagogical and ideological strands of distance learning. This novel combination attracted worldwide attention.

The OU slogan 'open as to people, open as to places, open as to methods and open as to ideas' encapsulates this. 'Open to people', means that there are no admission requirements for undergraduates. Today the Open University has 250,000 enrolled students. Yet despite its size it ranked 5th, one place above Oxford, in national assessments of teaching quality until the elite universities, which did not like this form of assessment, pleaded successfully with the authorities to stop it!

Note also that the Open University came top in the 2012 nation-wide assessment of students' satisfaction with their universities and has never come lower than third in this annual survey.

We conclude from this first, that you can deliver high-quality education to large numbers using technology and, second, that using media in education is an evolutionary process. The Open University has not changed its mission of openness to people, places, methods and ideas. However, between 1970 and 2010 the way that it expressed and implemented those values steadily evolved.

C.S. Lewis once wrote: "Humanity does not pass through phases as a train passes through stations: being alive, it has the privilege of always moving yet never leaving anything behind." That is an important principle in the use of media in education. A great strength of the Open University is that 'it is always moving on yet never leaving anything behind'. In the 1970s it revolutionised correspondence education and used broadcast TV and radio to fulfil its mission. Today it is the largest presence on iTunesU, with 60 million downloads of its material in the last five years, one sixth of them in China. The large majority of its 250,000 students engage in online learning leading to academic qualifications

3 Choice of Content

We shall now examine how easier access to learning content has contributed to the development of online learning. In the 1970s, at the same time that the Open University opened the great American educator Ernie Boyer, then Chancellor of the State University of New York, set up Empire State College with the aim of opening up the curriculum. It allowed students to work with mentors to invent their own courses of study, captured in its slogan 'my degree, my way'. With sound mentoring students could design credible programmes and courses for themselves. Thanks to the Internet and to the concept of Open Educational Resources the tools that students can use for this have expanded dramatically.

The notion of making academic content freely available for re-use and adaptation made news in the late 1990s when MIT started putting its lecturers course notes on the Web. UNESCO held a forum in 2002 to explore the implications of MIT's initiative for developing countries. The Forum coined the term Open Educational Resources (OER) and defined them as educational materials that may be freely accessed, re-used, modified and shared.

Ten years later UNESCO held a World Congress on which approved the Paris Declaration on OER. Its key recommendation is to encourage the open licensing of educational materials produced with public funds. There are signs that some governments are already taking the Paris Declaration and the economic benefits of OER seriously. For example, the province of British Columbia will now offer free, online open textbooks for the 40 most popular postsecondary courses.

4 Making Sense of MOOCs?

MOOCs have generated more media interest in the use of technology in higher education than any development since the Open University. The question for us today is how MOOCs will contribute to the development of online learning generally.

The term MOOC was invented in Canada in 2008 to describe an open online course at the University of Manitoba. The course, *Connectivism and Connective Knowledge,* was presented to 25 fee-paying students on campus and 2,300 other students from the general public who took the online class free of charge. The course was inspired by Ivan Illich's philosophy in his book *Deschooling Society* that an educational system should 'provide all who want to learn with access to available resources at any time in their lives; empower all who want to share what they know to find those who want to learn it from them; and, finally furnish all who want to present an issue to the public with the opportunity to make their challenge known'. In this spirit 'all the course content was available through RSS feeds, and learners could participate with their choice of tools: threaded discussions in Moodle, blog posts, Second Life and synchronous online meetings'. These courses were a logical development of the Open Educational Resources movement.

These early MOOCs, which are now called cMOOCs (for 'connecting' MOOCs), are very different from the next phase of MOOCs, called xMOOCs after edX, the MIT, Harvard and UC Berkeley consortium offers them. Those first xMOOCs had nothing to do with Illich's liberal educational philosophy and little relation to the pioneering cMOOC courses. However, MOOCs are now evolving rapidly, leading to some blending of the cMOOC and xMOOC approaches.

MOOCs are an interesting development in higher education for various reasons. A first is that elite universities, which have always restricted student access, are offering open enrolment. A second is that MOOCs have high dropout rates and very low pass rates. The figures may improve somewhat over time as the novelty wears off but will remain a serious problem. A third feature is that most universities offering MOOCs do not award credit for them, even where curricula of the MOOC and the equivalent course on campus are identical.

So where will MOOCs lead higher education? There are positives and negatives. One positive aspect is that the press coverage of MOOCs has created greater public awareness of open, distance and online learning. The downside is that few people complete MOOCs successfully and that even they do not get credit. Another positive is that we have a new pedagogy in higher education to augment or replace lecturing. But this requires MOOCs faculty to refine their online pedagogy as a mainstream teaching activity rather than a public relations sideline.

These upsides and downsides bring us back to the fundamental contradiction of MOOCs, which is the tension between offering online learning openly while recruiting regular students selectively. The key to getting a degree from elite institutions is to be admitted. It is more difficult to secure admission than to exit with a degree. For such institutions to adopt the opposite open-university principles of open admission and rigorous exit requirements would require a tremendous paradigm shift. We stress that this shift is an issue of mentality, not of media, technology or practicalities, because the open universities have shown that degree-credit programmes can be offered successfully to thousands of students using online methods.

Recently the authors have worked with two South African authors, Neil Butcher and Merridy Wilson-Strydom, to develop a Guide to Quality in Online Learning as an OER under a Creative Commons CC-BY-SA license [1]. The publisher, Academic Partnerships, specialises not in MOOCs but in helping universities to offer their regular degree programmes online for credit to larger numbers of students.

The Guide draws on examples from all over the world and aims to help the process of bringing online learning into the mainstream of higher education. We hope that MOOCs, despite their current contradictions, will contribute to that process.

Reference

1. Butcher, N., Wilson-Strydom, M.: A Guide to Quality in Online Learning (Uvalić-Trumbić, S., Daniel, J. (eds.)) Academic Partnerships, Dallas (2013)

Analysing the Impact of Built-In and External Social Tools in a MOOC on Educational Technologies

Carlos Alario-Hoyos[1], Mar Pérez-Sanagustín[1], Carlos Delgado-Kloos[1], Hugo A. Parada G.[1], Mario Muñoz-Organero[1], and Antonio Rodríguez-de-las-Heras[2]

[1] Department of Telematic Engineering, Universidad Carlos III de Madrid
Av. Universidad, 30, 28911, Leganés (Madrid), Spain
{calario,mmpsanag,cdk,hparada,munozm}@it.uc3m.es
[2] Faculty of Humanities, Communication and Documentation,
Universidad Carlos III de Madrid
Calle Madrid, 126, 28903, Getafe (Madrid), Spain
heras@hum.uc3m.es

Abstract. MOOCs have been a disruptive educational trend in the last months. Some MOOCs just replicate traditional teaching pedagogies, adding multimedia elements like video lectures. Others go beyond, trying to engage the massive number of participants by promoting discussions and relying on their contributions to the course. MOOC platforms usually provide some built-in social tools for this purpose, although instructors or participants may suggest others to foster discussions and crowdsourcing. This paper analyses the impact of two built-in (Q&A and forum) and three external social tools (Facebook, Twitter and MentorMob) in a MOOC on educational technologies. Most of the participants agreed on the importance of social tools to be in touch with their partners and share information related to the course, the forum being the one preferred. Furthermore, the lessons learned from the enactment of this MOOC employing social tools are summarized so that others may benefit from them.

Keywords: MOOCs, Social Tools, Educational Technologies, Crowdsourcing.

1 Introduction

MOOCs (Massive Open Online Courses) are considered one of the main educational trends in the last months [1, 2]. Initiatives like Coursera, edX, Udacity or MiríadaX are changing the ways we teach and learn, going beyond traditional online courses, and reaching thousands of learners worldwide [3]. The opportunity for free training through courses taught by experts from major Universities in a ubiquitous context seems very attractive for learners, opening up new opportunities for defining new pedagogies [4] and business models [5].

MOOCs are deployed in platforms that offer services for managing the massive amount of learners. The role of instructors in MOOCs is to design the initial contents, the assignments and the assessment activities that they later upload to these platforms. However, instructors play a secondary role during the enactment of MOOCs, compared

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 5–18, 2013.

to traditional online courses, since they cannot provide personalized support to the massive number of participants [6, 7]. The community of learners registered in MOOCs is expected to assist their partners, and to enrich the course with discussions and related contents as a way of crowdsourcing [8]; that is called "learner as teacher as learner" model [9]. The instructor becomes a "guide on the side" [10] clarifying only those key questions that drive the debate [11].

One of the main characteristics of current MOOCs is the high attrition rate among registered users [12], which can go up to 90-95% [1]. This is partially due to the free nature of most MOOCs, which attracts many observers that are not really interested in the contents delivered. However, many other users that are interested leave the course earlier than expected. Among their reasons, their difficulties to become self-learners and the lack of personalized support from instructors [8]. One approach to tackle this problem is to offer several social tools during the MOOC enactment in order to create a community of participants that provides support and advice to those with difficulties, connecting learners at an emotional and value level [13].

MOOC platforms normally include some built-in social tools like forums to centralize learners' contributions, discussions and queries; but also, instructors may suggest alternative tools, external to the platform. However, instead of introducing a pool of social tools randomly, it should be detected which ones are adequate to effectively build connections and collaboration among learners [7, 14]. Moreover, a proper selection of social tools can be a good mechanism to engage learners and promote their participation during the course, as outlined in recent guides for MOOCs design [15]. In summary, a proper selection of social tools can facilitate the community with the necessary support to advance in the course and may help reducing drop-outs from those learners that are interested in the course subject.

This paper proposes a deep analysis of how social tools are perceived and utilised by MOOC learners to shed some light on their selection process. Particularly, this analysis involves five social tools used throughout six weeks in a MOOC on educational technologies taught in Spanish and deployed in the platform MiríadaX by Telefónica Learning Services[1]. Two of the tools are built-in (a Q&A service and a forum), while three are external to the platform (Facebook[2], Twitter[3] and MentorMob[4]). The analysis on these tools includes their level of activity in order to detect which social tools are more actively employed, and the kind of information instructors and learners share through them.

The remaining of this paper proceeds as follows: section 2 presents the design of the MOOC on educational technologies and its deployment in the MiríadaX platform, including the built-in and external social tools chosen; section 3 overviews the enactment of the course, collecting information about learners' profiles and performance; section 4 analyses the level of activity in the five social tools and learners' perception on them; section 5 discusses the lessons learnt from the use of these social tools in the MOOC, with conclusions and future work in section 6.

[1] http://miriadax.net/web/educacion_digital_futuro
[2] https://facebook.com
[3] https://twitter.com
[4] http://mentormob.com

2 Design and Deployment of the MOOC

Five Professors and teaching assistants from the Universidad Carlos III de Madrid (Spain) participated in the design and deployment of the MOOC on educational technologies. The six weeks of this MOOC under analysis were structured into two modules covering two fields of knowledge: humanities and engineering. Thus, a wide range of learners could be reached no matter their backgrounds. Particularly, lectures on humanities (module 1) dealt with the concept of interaction and its application to the digital world, while those in engineering (module 2) were about the use of mobile devices in educational settings. Each module was taught during three consecutive weeks. An introductory module presenting the course context, structure, assessment and the social tools was released the day before the first module started.

2.1 Design of the MOOC

The MOOC was designed considering three different aspects in both modules: learning contents, assignments and assessment activities. Learning contents included 8-9 short videos of about ten minutes each week (24-27 videos per module) and supporting materials (i.e. the slides used in some videos); the videos contained weekly interviews with experts on the delivered topics. Assignments included additional reading material and a set of open questions in form of a video showing the opinion of both students and teachers at the campus. Finally, assessment activities covered formative assessment with multiple choice tests after each video lecture to reinforce the explained concepts; and summative assessment with one multiple choice test every week, and a peer review activity at the end of each module, where learners had to submit a work related to the contents explained in that module and later review some peers' works following a given rubric. The contents, assignments and assessment activities were available at the beginning of the week and remained open throughout the course, except for summative assessment activities, which were due at scheduled intervals.

2.2 Deployment of the MOOC

The MOOC was deployed in MiríadaX. This platform allows defining a course structure arranged in different modules, including multimedia resources, multiple choice tests and peer review activities. Videos were uploaded to Youtube and later embedded in MiríadaX as multimedia resources. The course could be followed from laptops, smartphones or tablets facilitating a ubiquitous participation and learning. The platform also provides features to send massive emails to the registered learners, and to publish announcements related to the course in a blog. Also, MiríadaX offers the following built-in social tools that were added to the MOOC:

- **Q&A.** This is a tool for learners to make questions about the enactment of the course, the contents of the modules or the platform. The instructors and other participants may answer the questions or vote them as relevant to gain visibility.

- **Forum.** This is a tool for learners to participate in discussions on selected course topics. The learners must maintain it, but instructors can define a thread structure and make comments on relevant learners' observations about course issues.

Three extra external social tools were selected for this MOOC:

- **Facebook.** Instructors can use this tool to send announcements, foster discussions and share additional multimedia resources with learners, who can contribute to the discussion and share new resources too.
- **Twitter.** Instructors can employ this tool to send short announcements, links to additional resources and quotes extracted from video lectures and learners, who can contribute disseminating the course to their followers.
- **MentorMob.** Instructors and learners can classify and share reading material and websites related to each of the modules through this tool.

Finally, two other external tools were selected in this MOOC: Storify[5] to arrange and share a collection of relevant tweets every week; and Google Drive[6] to deliver questionnaires about learners' profiles and degree of satisfaction with the course. These tools were only used for the instructors to collect and show the learners a summary of the activity in the social tools around the course.

3 Enactment of the MOOC

The two modules of this MOOC under analysis were enacted in February – March 2013. The course was announced three weeks before starting in Spanish and Latin American universities, social networks and press. The first day of the course there were 3105 registered participants. Nevertheless, the registry was never disabled and many learners joined later. All the contents were available since the week they were released to the end of the course, except for summative assessment activities, which were due in scheduled intervals. Even though many latecomers missed some of the first summative assessment activities, the instructors encouraged them not to leave the course since they could pass it by successfully accomplishing the remaining ones. The number of registered users after the six weeks was 5455.

3.1 Learners' Profiles and Motivation

Registered participants had to fill out a questionnaire to help instructors detect profiles and motivation as part of the introductory module. 3,362 learners submitted the answers to the questionnaire (44.6% men and 55.4% women). The range of ages was very varied with most participants between 25-35 years (37.5%) and 35-45 years (24.4%). Learners were located in 40 different countries (mainly Spain and Latin America), which represents a high impact of this MOOC reaching people through multiple frontiers, and at the same time a challenge to teach people with so many different cultural backgrounds. Top countries by number of participants were Spain

[5] http://storify.com
[6] http://drive.google.com

(59%), Colombia (10%), Mexico (8%) and Peru (6%). Though most learners had Spanish as their mother tongue (96%), there were important communities of native Portuguese speakers from Portugal and Brazil. It is important to stress that most of the participants combined learning through this MOOC with their regular jobs (57%) and studies (18%), thus demonstrating the potential of MOOCs for workers' training in their leisure time. Also, this MOOC exemplified the cost-effective opportunity for keeping active and updated those unemployed (22% of the registered participants).

The topics of this MOOC awaken interest mainly among qualified people: 40% of the learners took this course as part of their postgraduate studies, and 52% as part of their degree studies. Their backgrounds however were very different: 32% owned some technical background; 46% had some educational background; and 32% claimed some background on humanities. This represented a challenge for instructors, who had to adapt their explanations and vocabulary to a wide audience. Despite the lower number of people with technical background, the second module (i.e. the use of mobile devices in education) was a priori of interest for double those of the first module (i.e. the concept of interaction and its application to the digital world). Finally, it is convenient to note that 61% of the learners had experience teaching and 30% had even recorded educational videos.

3.2 Learners' Performance

From the 5455 registered participants only 81.6% of them started the presentation module, and 64.7% the first week of the first module. This suggests that more than a third of the registered users had no real interest in the actual course material, this being aligned with the percentages highlighted by other researchers [16]. Significantly, there was another sharp drop after the first week, and only 40.7% of the registered learners watched any of the videos scheduled for that week. According to their behaviour, those users that leave the course before the second week (i.e. they enroll in the course but just look at a few contents at the most) are classified in [17] as **lurkers** (either no shows and observers in [18]) and represented 59.3% of the registered users in this MOOC on educational technologies (see Figure 1).

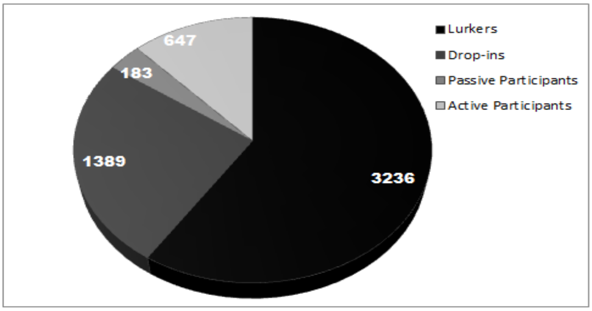


Fig. 1. Classification of the users registered in the MOOC on educational technologies according to their behavior. The description of the type of participants is adopted from [17].

After the second week, the number of participants leaving the course progressively decreased: 31.5% entering the third week; 25.6% the fourth week; 22.2% the fifth week; and 19.7% the sixth week. 830 (15.2%) learners answered the sixth week final test. Users that are initially interested but leave the course as the weeks go by can be classified as **drop-ins** [17, 18], and represented a 25.5% in this MOOC (see Figure 1). Their reasons to stop participating may be interest only in selected topics, disillusionment with the course, lack of support from instructors, lack of time to complete the assignments, or bad experiences with the MOOC platform. Here, learners reported the tight schedule and some problems with MiríadaX as the main shortcomings.

Those users that reach the end of the course can be classified in two groups: **passive participants** and **active participants** [17, 18]. Passive participants are those that consume video lectures and maybe take some tests, but that are reluctant to participate in discussions or complicated activities. On the contrary, active participants not only consume contents, but also contribute to the discussions using social tools, and regularly take part in more elaborated activities such as peer-to-peer (P2P) assessments. In this MOOC we classified as active participants those that claimed to have worked more than two hours per week in the course (the time to see all the video lectures and complete the formative and summative multiple choice tests was estimated in about two hours per week). According to this criterion, 11.9% of the registered users were classified as active participants and 3.4% as passive participants (see Figure 1). These data were collected from a survey delivered after the end of the second module. The coherence of these figures was checked comparing them with the number of learners that completed the P2P assessment in the second module and with the use of social tools (see section 4).

Figure 2 shows the number of learners that carried out the main milestones of the course in chronological order: the surveys about learners' profiles and degree of satisfaction with the course, the weekly multiple choice tests employed for summative assessment and the P2P activities also employed for summative assessment. It is noteworthy that while tests and P2P activities were due in scheduled timings, the surveys were open throughout the course. Differences between active and passive participants are denoted comparing those that carried out the tests and the P2P activities. The P2P activity in the second module received more learners than the same activity in the first module. A possible explanation for this is the initial confusion in the procedure for submitting and reviewing peers' work, as noted from dozens of learners' comments in this line within social tools.

4 Analysis of the Social Tools Employed in the MOOC

As explained in section 2.2, five social tools were employed for different purposes as part of this MOOC on educational technologies: two built-in MiríadaX tools (Q&A and forum) and three external tools (Facebook, Twitter and MentorMob). The analysis of the impact of these social tools was carried out from two different points of view: learners' perspective indicating the utilization of the social tools, and tools perspective, collecting quantitative data from their actual use.

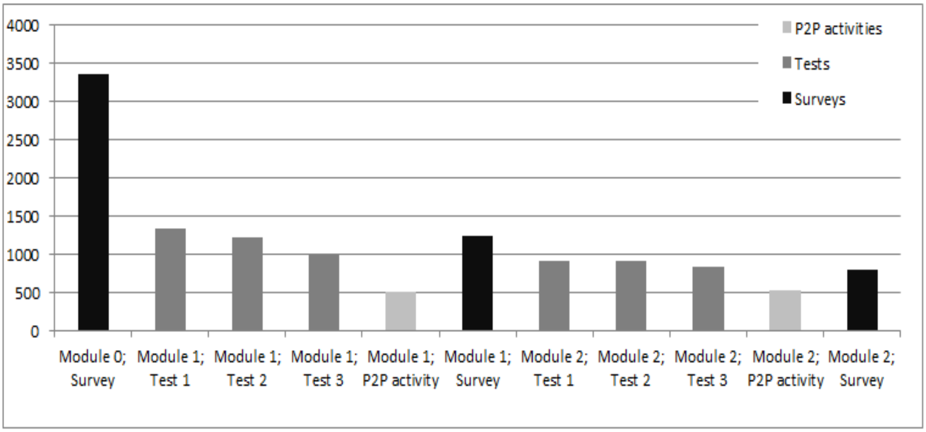


Fig. 2. Number of learners that carried out the main MOOC milestones

4.1 Analysis from Learners' Perspective

Learners were asked whether they intended to participate in the social tools of this MOOC twice during the course: before starting the first module, and before starting the second module. Most learners were willing to participate, and only less than 10% of them had clear from the very beginning that they would not contribute to the discussions.

At the end of both modules, learners' assessed their own participation in each social tool using a six-point Likert scale where one meant, "*I did not use this tool*", and six, "*I used this tool very actively*". Table 1 summarises the results including also the number of answers collected from the surveys. These data indicate that the built-in MiríadaX forum was the social tool that had a higher impact in this MOOC, followed by Q&A, which was very actively used by about 5% of the participants. From the external social tools, Facebook was the one that received a higher interest from learners. It is important to note the growing use of the forum, Facebook and Twitter during the second module. There are two possible explanations for this behaviour: because passive participants are given to leave MOOCs earlier; or because some passive participants may become active after a while benefitting from the cluster of contributions generated by their peers. On the other hand, Q&A kept a similar level of activity while that of the MentorMob slightly decreased in the second module, without catching learners' attention at the level expected by instructors. Learners confirmed the overall increase of participation in the second module: 30.3% of them reported having increased their contributions in the social tools versus 24.8% who decreased them.

To complement this analysis, learners were asked at the end of each module whether they believed social tools were a good mechanism to be in touch with the other participants and share additional information to that provided by the instructors. About 65% of learners reacted positively to this assertion and only 11% of them expressed disagreements, with barely any differences between modules.

Table 1. Learners' perception of their use of the social tools after modules 1 (M1) and 2 (M2)

	Built-in				External					
	Q&A		**Forum**		**Facebook**		**Twitter**		**MentorMob**	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
1 (I did not use this tool)	47.8%	47.1%	45.1%	39%	59.9%	54%	71.6%	65.4%	63.1%	63.3%
2	21%	21.1%	23.6%	22.3%	17.3%	17.3%	10%	14.1%	15%	18.9%
3	14.6%	14.1%	16.7%	21.3%	11.1%	13.7%	8%	9.1%	10.9%	9.7%
4	6.1%	6.8%	8.3%	9.6%	5.7%	7.6%	4.5%	6.4%	5.2%	4.6%
5	5.5%	5.4%	3.8%	5.5%	3.8%	4.2%	3.2%	2.7%	3.5%	2.5%
6 (I used this tool very actively)	5%	5.5%	2.5%	2.4%	2.1%	3.2%	2.6%	2.3%	2.1%	1%
Number of answers	1190	760	1188	764	1176	761	1173	749	1164	755

4.2 Analysis from Tools Perspective

Q&A had a moderate impact throughout this MOOC. It was frequently consulted by the learners, who mainly inquired about the course assignments, assessment activities and certification, occasionally setting out also questions related to the concepts explained in the theoretical video lectures. Also, it was the entry point for complaints about technical problems in the platform MiríadaX. The instructors regularly replied these questions, but also many learners further contributed with answers or stressing the same questions to gain visibility. At the end of the course there were 270 questions posted, and 464 contributions either questions or answers from 273 different learners. These questions and answers were visualized 6,485 times.

The **forum** was the social tool with a higher impact during the enactment of this MOOC. It was employed for long discussions about the course topics and to exchange common interests within the community of participants. 730 different learners contributed to the forum, posting 2,382 messages in 721 threads. Instructors played a secondary role adding another 138 messages to the most controversial threads. Interestingly, several learners boosted the debates. One of the learners, for instance, posted 155 messages and another participant 82; there were 42 learners with 10 or more contributions and 155 with five or more. The messages posted in the built-in forum were seen 20,901 times.

Facebook produced a moderate impact on the learners of this MOOC on educational technologies. Figure 3 represents on top the number of unique users that talked about the course website on Facebook per day (i.e. liked the page, posted a comment, liked a comment, or wrote about the site in another page). The bottom of Figure 3 extracts the distribution of new likes per day, which reached a total of 1,297, following a similar distribution worldwide to that of students' profiles in section 3.1 (e.g. 46% from Spain, 12% from Colombia, 11% from Mexico, 8% from Peru). Both graphs reflect an initial excitement and a progressive drop as it also happened with learners' performance in section 3.2. Interestingly, peaks on likes correspond to the start of new weeks, when instructors posted messages on Facebook with instructions about new available materials and deadlines. Other peaks that can be seen, but only in users talking about the course (Figure 3 top), are a consequence of learners' replies to instructors' open questions aimed at fostering the debate. Instructors published 28 messages on Facebook, and received 529 comments from 275 different learners (and 874 likes on these messages and comments). 44,236 users visited this site considering unique users per day according to Facebook analytics.

Twitter had a moderate impact on this MOOC, reaching 815 followers. It was mainly employed to send short announcements by the instructors, who posted 97 tweets (21 retweeted from learners and 19 automatically generated from Facebook entries). The MOOC hashtags facilitated the easy aggregation of Twitter messages produced by participants. Learners shared their opinions in the form of tweets, but also linked videos and other resources related to the course topics. 165 different learners mentioned the course hashtags, although a few of them were particularly active with up to 26 posts. Figure 4 shows a distribution of the times the MOOC hashtags were mentioned per day. It is noteworthy that this distribution follows a similar structure compared to that of Facebook (Figure 3), since announcements and open questions were sent to both social tools at the same time.

MentorMob did not reach the expected impact on this MOOC as a tool to share contents. Instructors created three lists with additional material: one for the presentation module with three personal sites, one for the first module with three readings related to the interaction topic, and one for the second module with 17 additional resources about the use of mobile devices in education. Only 43 new contributions in total were received from 32 different learners. It was noted that students preferred other social tools like the forum or Facebook to share resources. The consumption of extra material was however significant with 20,590 visualizations among all the resources provided by the instructors in MentorMob and those crowdsourced by the learners using this social tool.

Table 2 summarizes the contributions from learners in the form of posts in the different social tools indicating also the number of posts from the most active learners, who turned out to be different in each social tool. The built-in MiríadaX forum was the tool with a higher popularity among the participants on this MOOC on educational technologies, followed by Facebook. These results are consistent with those presented in section 4.1 about learners' perception of their use of social tools, although in that section both production and consumption were implicitly analysed, while here only production is taken into account.

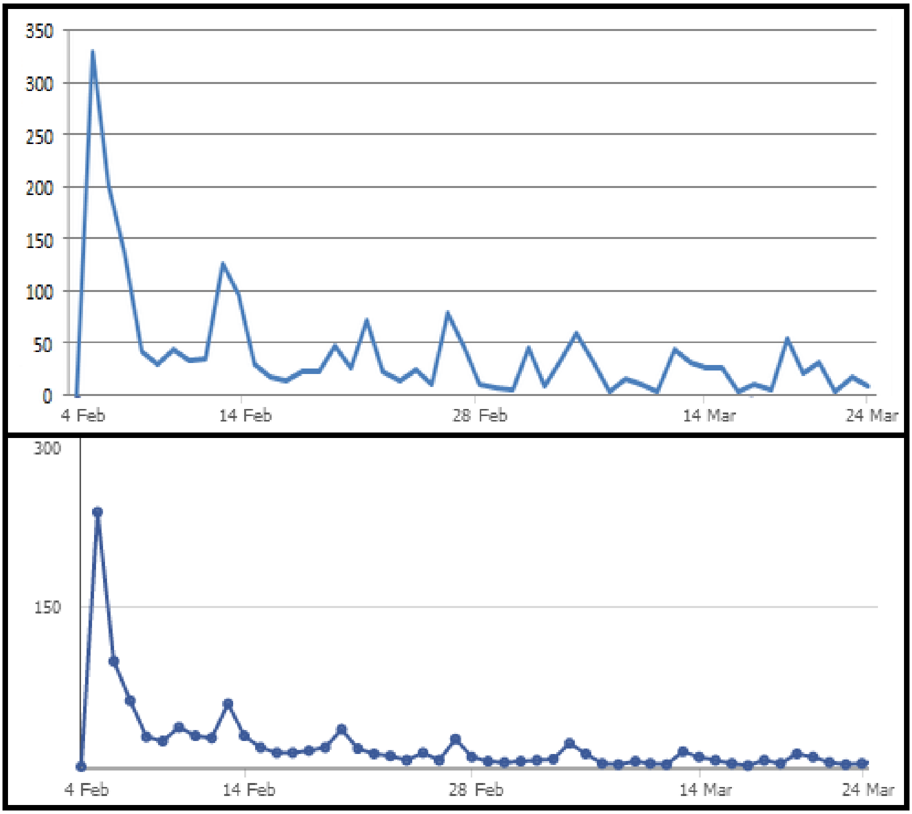


Fig. 3. Facebook impact on MOOC learners: graph on top represents the unique users that talked about this MOOC on Facebook per day; graph on bottom represents the unique users that liked the Facebook site about this MOOC per day. Source: Facebook.

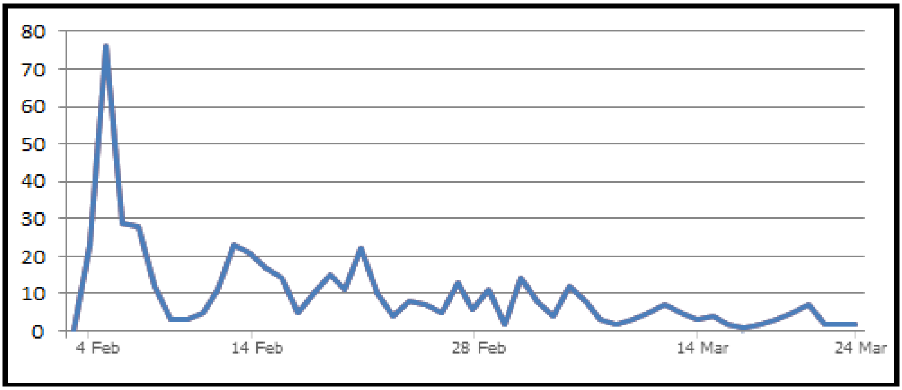


Fig. 4. Twitter impact on MOOC learners: tweets related to the hashtags of this MOOC per day. Source: Twitonomy and Topsy.

Table 2. Contributions from learners in the different social tools

	Built-in		External		
	Q&A	**Forum**	**Facebook**	**Twitter**	**MentorMob**
Number of learners that posted on this social tool	273	730	275	165	32
Number of posts	464	2,382	529	504	43
Number of posts by the most active learner	7	155	12	26	6

5 Discussion about the Use of Social Tools

The analysis of the impact of social tools in this MOOC suggests that built-in tools, and particularly the forum, are preferred by the learners to discuss and contribute to the MOOC. One possible explanation to this finding is that the centralization of tools and contents in one single platform allows participants to work in learning and assessment activities without exposing their personal lives as usually occurs when employing external social tools like Facebook or Twitter. This explanation matches with the results obtained in the study by Kop et al. [7], where people employed more the forums than the Facebook groups for privacy and personal security reasons and for a sense of trust and feeling comfortable and confident to be able to participate.

It is important to take into account that not all the tools cover the same purposes (i.e. while some social tools may be useful for short announcements or posing questions, others are specific for discussions). Further, learners may be used to several social tools in their daily lives, promoting their utilization when working in MOOCs. As an example, despite the low number of learners contributing in Twitter, their average number of posts was higher than the one in Facebook or Q&A, as it was also higher the number of posts of the most active learner. Therefore, it seems appropriate to offer multiple options of social tools when designing and enacting MOOCs in order to cover the wide range of expected learners. This is aligned with the conclusions extracted in [7], where authors point out the importance of different contexts in open courses to get different levels of participation and outcomes.

However, offering multiple social tools has a counterpart, since it also means decentralizing discussion threads and increasing the number of places that learners may need to check to follow their peers' contributions. Thus, it is important that both instructors and participants learn how to filter the massive amount of information that can be crowdsourced in a MOOC. MOOC platforms may include some services to facilitate this filtering (e.g. rankings for most voted posts); although external social tools does not always provide indicators of the quality of the contributions that help learners discard those with a lower relevance. Particularly, we received complaints of information overload diluted in several spaces (forum, Facebook, Twitter…) from some participants. One of them even pointed out that too much information may hinder the learning process. Thus, it would be recommendable to give advice to learners about how to filter information at the beginning of the MOOC. Moreover, best practices about how to use forums and social tools in a clean and ordered way should be taught and reminded periodically (e.g. not creating unnecessary threads, reading first before writing, etc.). Even though we provided documentation for this purpose at the beginning of the MOOC, the guidelines were not followed by some of the users, who created a bit of mess favoured by the massive context, especially in the forum.

Information overload together with a lack of appropriate searching mechanisms complicated that instructors could solve key questions related to the course topics and support students in advancing on the course. Results show that the students were who assumed the role of mentors, giving advice to those participants with higher problems to follow the course, and offering themselves for queries related to the assessment activities. This happened especially in Q&A, in the forum and in Facebook during the enactment of this MOOC. This supports the conclusions extracted in [7], which stress on building MOOCs based on the learner-in-dialogue model inspired in the conversational framework by Laurillard [19] and on the co-creation of the MOOC environment, with activities for reinforcing orientation for learners, coaching, mentoring and practices of peer facilitation. The role of volunteer mentor demands special features to those playing it, normally: advanced knowledge in the delivery subjects, engagement with the course structure and activities, flexible time to work in the course and attitudes to help the others without financial reward in exchange. From the thousands of people registered in a MOOC it is likely to find a few of them that may voluntarily play the role of mentor, as it was the case here. All this matches with the conclusions in [14], which highlight the high number of student to student interactions that normally occur in MOOCs, to fill the gap regarding student to instructor interactions.

Nonetheless, students do not always employ social tools in MOOCs to critically contribute to the discussions or to help learners in trouble. Their open nature makes that everyone can register and post comments, and so there is a risk that people with negative intentions try to divert attention from learning tasks. For instance, we detected a few users that registered just to post complaints, most times about unrelated subjects (e.g. political nature). Others tried to advertise their products, looking for customers in the mass of participants. Finally, a couple of users made public the answers to some of the multiple choice tests in the forum when others still had time to complete them. Of course, assessment is a very important weakness on current MOOCs that is under research with strategies like proctored exams [20]. However, our experience shows that a few scatterbrained or bad intentioned people may disturb the work of many others willing to learn. To fight against those undermining the learning process, MOOC platforms and external social tools should include options to unregister those users with a bad behaviour and to mark comments as inappropriate when proceeds. However, it is still an open challenge to see how to make these unregistering options open to the community of participants, since instructors cannot face the huge amount of information generated.

Finally, it is noteworthy that participants may decide to employ other social tools, apart from those selected by the instructors. For instance, in the MOOC on educational technologies one of the participants decided to share his work in a P2P activity in Google Documents posting later the URL in the forum to be assessed by his partners. The motivation for this fact was that the deadline for submitting the P2P works through the platform MiríadaX (as stated by the instructors), had expired and so, the corresponding submission feature was disabled. This is just one example, but there is a great potential in the community of learners for the selection of alternative tools for their communication and discussion within MOOCs.

6 Conclusions and Future Work

Current research points out that the community of learners registered in MOOCs should be responsible for assisting those peers with problems, who cannot find support from instructors due to the massive condition of these courses. Thus, a proper selection of social tools is needed to generate in the participants a sense of "place" and proximity that facilitates answering queries, fostering discussions, and contributing with new resources to the course.

In this paper, we analysed five social tools employed in a MOOC on educational technologies. The forum has revealed as the tool preferred by the learners for social interactions and discussions followed by Facebook and Q&A, although the latter was mainly employed for specific questions about course procedures rather than for deep discussions related to course topics.

One important lesson learnt from this analysis is the trade-off between providing multiple social tools to satisfy most learners' needs during the MOOC enactment, and the information overload that the use of all these tools at the same time causes. This overload may hinder the real learning process if participants get lost in all their partners' posts without properly filtering those that are relevant.

However, this study also poses a new set of open questions: 1) How to design MOOCs to foster and scaffold participation of novel learners while promoting the mentoring role of experts? 2) How to provide tools for supporting MOOC participants' self-regulation to deal with those trying to undermine the learning process? Considering that the forum and Facebook were the tools that received more contributions, 3) Would the activity of participants in social tools increase if we could automatically integrate those Facebook posts related with the forum open threads? These challenges open new research venues for other researchers involved in MOOCs.

Apart from dealing with these open questions, future work will analyse other social tools in MOOCs of a different duration. A different research line will look for mechanisms to engage more MOOC learners to contribute to the discussions in social tools. Finally, the connections between the actual use of social tools and the scores obtained by the learners in MOOCs will be researched.

Acknowledgements. This work has been funded by the Spanish Ministry of Economy and Competitiveness Project TIN2011-28308-C03-01, the Regional Government of Madrid project S2009/TIC-1650, and the postdoctoral fellowship Alianza 4 Universidades.

References

1. Daniel, J.: Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. Technical report, Korea National Open University (2012)
2. Hyman, P.: In the Year of Disruptive Education. Communications of the ACM 55(12), 20–22 (2012)

3. Cooper, S., Sahami, M.: Reflections on Stanford's MOOCs. Communications of the ACM 56(2), 28–30 (2013)
4. Martin, F.G.: Will Massive Open Online Courses Change How We Teach? Communications of the ACM 55(8), 26–28 (2012)
5. Kolowich, S.: How will MOOCs make Money? Inside Higher Ed. (2012), http://www.insidehighered.com/news/2012/06/11/experts-speculate-possible-business-models-mooc-providers
6. Downes, S.: The Role of the Educator. Huffington Post Education (2010), http://huffingtonpost.com/stephen-downes/the-role-of-the-educator_b_790937.html
7. Kop, R., Fournier, H., Mak, J.S.F.: A pedagogy of Abundance or a Pedagogy to Support Human Beings? Participant Support on Massive Open Online Courses. International Review of Research in Open and Distance Learning 12(7), 74–93 (2011)
8. Mackness, J., Mak, S., Williams, R.: The Ideals and Reality of Participating in a MOOC. In: Dirckinck-Holmfeld, L., Hodgson, V., Jones, C., de Laat, M., McConnell, D., Ryberg, T. (eds.) 7th Int. Conference on Networked Learning, pp. 266–274 (2010)
9. Siemens, G.: Knowing knowledge. Lulu Press, Vancouver (2006)
10. Masters, K.: A Brief Guide to Understanding MOOCs. The Internet Journal of Medical Education 1(2) (2011)
11. Severance, C.: Teaching the World: Daphne Koller and Coursera. Computers 45(8), 8–9 (2012)
12. Clow, D.: MOOCs and the funnel of participation. In: 3rd Conference on Learning Analytics and Knowledge, pp. 185–189. ACM, New York (2013)
13. Shedroff, N.: Experience design 1.1: A manifesto for the design of experiences. Experience Design Books (2009)
14. MoocGuide, 4. Designing a MOOC using social media tools, http://moocguide.wikispaces.com/4.+Designing+a+MOOC+using+social+media+tools
15. Khalil, H., Ebner, M.: Interaction Possibilities in MOOCs – How Do They Actually Happen? In: Proceedings of the 3rd International Conference on Higher Education Development (accepted, 2013)
16. Belanger, Y., Thornton, J.: Bioelectricity: A Quantitative Approach, Duke University's First MOOC. Technical Report, Duke University, NC, USA (2013)
17. Hill, P.: Emerging Student Patterns in MOOCs: A Graphical View (2013), http://mfeldstein.com/emerging_student_patterns_in_moocs_graphical_view
18. Hill, P.: Emerging Student Patterns in MOOCs: A (Revised) Graphical View (2013), http://mfeldstein.com/emerging-student-patterns-in-moocs-a-revised-graphical-view
19. Laurillard, D.: Re-Thinking University Teaching: A Framework for the Effective Use of Educational Technologies. Routeledge, New York (1993)
20. Young, J.R.: Coursera Announces Details for Selling Certificates and Verifying Identities. The Chronicle of Higher Education (2013), http://chronicle.com/blogs/wiredcampus/coursera-announces-details-for-selling-certificates-and-verifying-identities/41519

Using a 3D Online Game to Assess Students' Foreign Language Acquisition and Communicative Competence

Anke Berns[1], Manuel Palomo-Duarte[2], Juan Manuel Dodero[2], and Concepción Valero-Franco[3]

[1] Facultad de Filosofía y Letras
[2] Escuela Superior de Ingeniería
[3] Facultad de Ciencias de la Educación
Universidad de Cádiz, Spain

Abstract. Over the past decade there has been an increasing attempt to explore the potential of computer games in order to engage students' towards foreign language learning. However, literature review has shown that there are still few attempts to provide empiric evidence of the educational potential of computer games, whereas the present study aims to address this lack. The purpose of our paper is to explore the possibilities of using a highly interactive 3D online game, we previously designed according to our student needs, in order to reinforce their foreign language acquisition and communicative competence. The target groups are students of a German foreign language course from the A1 level (CEFR). For our project we designed a 3D online-game that is based on a role-play, in which students need to cooperate in order to complete the game successfully. The game is built upon the OpenSim platform, and cooperation is conducted through game chats. Game chat interactions are being registered in a log file that is later processed. Furthermore, by analyzing different indicators we are able to obtain initial evidences to assess students' proficiency regarding their communicative competence in the target language. We include some examples from a pilot study we did with students of a German foreign language course (A1).

Keywords: Cooperative learning, virtual worlds, computer-supported language learning.

1 Introduction

The current paper presents the results of an ongoing project which aims to explore the possibilities of enhancing students' foreign language acquisition and communicative competence through the use of a 3D game-based environment. The starting point of our research is the increasing need to provide students with those learning tools which allow them to foster their foreign language skills beyond the classroom. The reason is that our students are usually enrolled in very large size German language courses, whereas they often lack opportunities to use

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 19–31, 2013.

and interact with other speakers in the target language. Hence, language practice outside the classroom becomes extremely important in order to complement face-to-face teaching and thus to guarantee that students reach the competences they are expected to have at the end of the term.

With the increasing rise of the Information and Communications Technologies, teachers and academic institutions often integrate blended teaching practices based on the use of virtual learning environments (VLEs). The latter aim to provide teacher and students with access to learning contents, services and applications anytime, any place and at any pace. Most widely used LMSs offer a huge variety of learning tools (i.e. wikis, discussion forums, blogs, chats and messaging) which facilitate interaction beyond the classroom ([3,1]). There has been furthermore an increasing attempt to explore the educational potential of computer games in the area of foreign language learning ([16,23,7]). This has been done both through adaptations of a variety of commercial games such as World of Warcraft, The Sims, Wonderland, Lineage II or Ragnarok Online (see amongst others [21,25,22]) as well as through the specific design of educational games, also called serious games ([4,7]).

In recent years many research studies have made claims on the benefits of integrating Multi-User Virtual Environments ([10]) and, especially, computer games in the area of foreign language teaching and education, stressing their motivational as well as educational potential (see amongst others [30,21,1]). Nonetheless, a number of researchers have underlined that there are still very few empiric studies which have directly investigated the effects of computer games on students' foreign language learning ([23,7,21]). In addition to this, literature review has shown that computer games are often employed to address very specific aspects such as vocabulary and grammar training ([8,9,5]). This is often done through drill-based activities rather than through authentic and synchronous peer-interaction in the target language.

Despite the still widespread use of drill-based computer games, there has been recently an increasing trend to explore the potential of computer games to enhance foreign language learners' interaction and communicative competence. As outlined by Garcia-Montero [12] in order to acquire communicative competence one needs to use the target language by interacting as much as possible with other speakers.

The aforementioned aspects, together with the increasing need to engage students in self-directed learning beyond the classroom, may explain the recent interest in exploring the pedagogical as well as motivational potential of massive-multiplayer-online games (MMPOG) as well as synthetic immersive environments (SIEs). Both are based on the interaction as well as collaboration with other players, providing thus multiple opportunities for authentic synchronous interaction, in and with the target language (recently [22,19,6,26,7,21]). A deeper look at some of the key-principles of foreign language acquisition may shed some further light on the learning potential of those computer games, which provide highly interactive environments.

2 Theoretical Background

2.1 Computer Games and Foreign Language Acquisition

Taking into consideration that the development of a second language is almost wholly dependent on the amount of comprehensible input that a learner receives ([14,31]), several researchers have underlined the importance of providing foreign language learners with those kinds of interactions that facilitate the intake of comprehensible input ([15,20,11,13,2]), stimulate the production of language output ([29]), encourages negotiation of meaning ([20]) and, last but not least, facilitates noticing ([24]). According to Swain's output hypothesis, "output production" is seen not only to enhance learners' fluency in the target language, but also to make learners "conscious of their deficits (...)". By doing so learners are encouraged to reflect on their language ([27]) and hence to improve their language skills. As stressed by Swain [29] "pushed output" is an important factor to focus learners' attention on aspects such as form and feedback. The positive and negative feedback learners get during their interaction with others is seen as a key-issue for foreign language acquisition, since it provides learners with the possibility to modify their language output and therefore to develop their competences in the target language ([24,15,27]). The above mentioned aspects make clear why versatile interaction with other speakers of the target language, independently if these are native or non-native speakers, becomes a key issue for a successful foreign language acquisition ([3]).

Bearing in mind the aforementioned aspects the current project aims to explore the potential of a 3D online game (*The hidden room*) to reinforce students' communicative competence through versatile interaction. The game has been designed by us within a virtual platform, called OpenSim, which allowed us to build several islands (virtual spaces) and game levels. Taking into consideration that our target students are from the A1 level (CEFR) we considered it extremely important to design a game that could easily be played by our students, providing them with very clear tasks and goals. Furthermore, in order to design learning tasks which meet our students' language proficiency as well as needs regarding language practice beyond the classroom, we -as a teacher and game-designer-needed equally to control and administrate the game environment.

This is why we believe that neither virtual worlds (VWs) such as Second Life, Active Worlds etc. nor MMPOGs are the most appropriate VLEs for our target group students. Whilst VWs are usually very open spaces which often lack of clear goals ([1]), MMPOGs generally require both, certain game-expertise and confidence with the target language ([21]). By designing our own game-environment the current study aims to give some empiric evidence of the educational potential games may have for A1 level students' foreign language learning, when designed according to their needs. Our study focuses therefore mainly on the following questions:

– How does participation in a cooperative online game affect students' interaction and communicative competence in the target language?

– Which factors influence students' interaction and communication with others while playing the game?
– Does the online game enhance students' fluency and accuracy in the target language?

2.2 Instructional Framework

The instructional framework used for the study and game design is based on the principles of cooperative learning (CL). The CL strategy is based on grouping students in small learning teams to work in cooperation with each other in order to solve a problem, or to perform a task presented by the instructor. Since its rise in the early 1970s many researchers have stressed that learners, who work together, acquire more language and social skills than those who study the same content under individualistic circumstances ([32]). Researchers have equally stressed that to ensure the optimum of opportunities for interaction CL should take place in small groups, since this usually works best. In order to guarantee ideal conditions for CL Stenlev [28] and Oxford [18] underline the importance of the following principles: *Simultaneous interaction, equal participation, positive interdependence, individual accountability, cognitive development as well as social development.*

3 Methodology and Experimental Environment

As outlined before, the primary purpose of the current study is to analyse the impact of a 3D online-game (*The hidden room*) on A1 level students' foreign language acquisition and communicative competence. In order to provide learners with a highly interactive learning environment we designed a 3-D online game which requires students' cooperation and thus interaction in the target language. By hosting the game in our own server we were able to store and analyze students' interactions within the text-chats. The game aims to give students the opportunity to practise one of the most difficult grammar items of German language, which refers to the use of changing prepositions. Depending on the context in which these prepositions (under, over, behind, on the right side of, at the back, at the front, etc.) are being employed, they need to be used in combination with different cases (accusative or dative) and articles (masculine, feminine or neuter). In order to familiarize our Spanish students with its correct use, we designed a game-application, based on two different levels: A first level, which aims to introduce the use of local prepositions through a game-based activity, called *Memory*, and a second level, called *The hidden room*, which aims to reinforce the previously introduced prepositions through a role-play activity. Whilst the first level can be performed solely individually, the second level needs to be performed cooperatively (see Figs. 1 and 2). In the following we will offer a more detailed description of each game level and the activities designed for it. Level one consists of a room with four different walls, each one of which is divided into two different but connected panels.

By clicking on the left panel several situations are being displayed by means of pictures (the hoover is on the right side of the table, the pizza is under the table, etc.). By clicking on the right panel the same situations are then being displayed by means of texts and audio-files. The students' task consists in exploring and matching the correct pairs. Once the students have matched a correct pair they get a score, that is immediately registered and displayed on the screen. In order to make the game more challenging a score-system, a time limit, different sound effects as well as feedback information (*Well done. Go ahead! Please repeat the activity! or You need to get 48 scores. Please try it again!*) have been included. All of them aim to provide players both, with individual feedback on their task performance and on how to perform the game successfully. Once the students have passed successfully the first level, they are allowed to proceed to the second game level. Unlike level one, level two aims to consolidate the previously (in level 1) acquired knowledge through a role-play activity. The second level is therefore based on a room which contains a number of different objects: clothes, furniture, food, beverages, toys and electronics. These are placed all over the room and need to be put in their correct position reconstructing the original room order. As the activity requires the cooperation of two players, we designed two different game tools: a controller and a viewer with different views of the room. At the beginning of the game the teacher or administrator provides each player with a different tool. In order to perform level 2 cooperatively player A gets a "controller" that enables her to move the objects within the room (see Figure 1) and player B gets a viewer with several views of the room (see Figure 2). Furthermore, in order to reconstruct jointly the correct position of each object, players A and B need to communicate only by using the text chat.

Fig. 1. View of player A with the controller tool

Fig. 2. View of player B. Picture shows the correct object positions.

4 Experimental Setting

In this section we comment the methodology for selecting the participants, the data collected (both from the platform and from pre- and post-tests) and its analysis. Further information can be obtained on the website[1] for this experiment. The virtual learning environment used to deploy the game-based learning experience is the OpenSim 3D virtual world (http://opensimulator.org/) open-source application server, through which a number of people and robots can interact according to computer programmed interaction rules. The 3D server is augmented with a back-end service and a set of tools that enable to log the data produced by interaction of avatars in the virtual world. The back-end service delivers data logs in a format set up for further data analysis through statistical and machine learning software tools, which can facilitate the observation of the virtual game-environment based on traces left by users in their collaborative activities ([17]).

4.1 Study Participants and Procedure

The observation of our students during more than 5 months allowed us to conclude that there are differences with regard to students' willingness to cooperate and interact with others in the target language. It was conspicuous that in general those students who used to interact more in face-to-face teaching obtained better learning results. Without the intention to analyze the games' appropriateness in terms of learning compared to more traditional method (as a control group), we here lean on our previous observations to make a preliminary analysis of the games' impact on students' willingness to interact. We therefore selected for our pilot study 12 students of whom 6 students were amongst the most active students and 6 amongst the less active ones. We then considered different categories in which we placed the more active students (category 1), the less active students (category 2) and a combination of both (category 3). Hence we created 6 teams (1A, 1B, 2A, 2B, 3A and 3B), two for each category (1A and 1B in category 1, 2A and 2B in category 2 and 3A and 3B in category 3). Even though we are aware that the sample size is extremely small to obtain final results, it represents an important first step within our ongoing study, which need to be done with a larger amount of participants. Only in this way we are able to investigate how the composition of different teams and the kind of feedback affects foreign language acquisition in non face-to-face settings.

4.2 Data Collection and Analysis

Two types of data were collected: 1) transcripts of students' produced chat dialogues were registered by the OpenSim platform and permit us to analyse students' interaction and use of the target language and 2) a pre- and post-test to

[1] http://code.google.com/p/daifceale/wiki/April_2013_ECTEL_2013

analyse the impact on students' foreign language acquisition and communicative competence.

In Table 1, we can see the number of turns (considering a turn when a student stops talking and the other starts), phrases (number of phrases used in the activity), single words (single-word phrases), total words (total amount of written words) and German-only total words (total amount of German written words).

Table 1. Number of words and number of phrases and turns in text-based chat

Category	1		2		3	
Team	1A	1B	2A	2B	3A	3B
Number of turns	88	38	77	75	43	57
Phrases	171	68	134	146	115	102
Single words	92	45	60	77	46	65
Total words	456	220	288	378	234	322
German only total words	349	220	281	367	231	317

Table 2 shows the results of students' interaction while playing the game. During their interaction students use a variety of discourse functions such as greetings, feedback, clarification requests, confirmation checks, paraphrasing, self-correction, wh-question and exclamations which permit them to advance jointly within the game and thus to complete the game task successfully.

Table 2. Discourse functions of clauses in the written interaction during the game-play

Category	1		2		3	
Team	1A	1B	2A	2B	3A	3B
Greetings	4	1	4	1	0	2
Feedback	24	15	38	23	37	19
Clarification requests	5	2	11	7	0	3
Confirmation checks	9	6	4	3	5	7
Paraphrasing	5	0	4	0	3	4
Self-corrections	2	0	3	2	1	0
Wh-questions	17	1	4	9	0	7
Exclamations	10	2	9	3	1	3

Additionally, Table 3 gives some further inside view with regard to students' interaction and use of the target language while playing the game and communicating within the text-chat. The results show that students seem to focus more on performing the game task successfully rather than on using the target language accurately. This may explain the relatively high percentage of grammar and lexical mistakes we can observe within the text-chat. Nonetheless, it is noteworthy that even though all participating players were from the A1 level of german language, almost all of them used solely the target language to communicate with the game-partner.

Table 3. Linguistic features of students foreign language production via text chat during the online game

Category	1		2		3	
Team	1A	1B	2A	2B	3A	3B
Grammar mistakes	28	14	29	18	25	35
Lexical mistakes (spelling)	43	30	53	76	28	26
Use of native language words	2	0	3	11	0	1
Total of words	456	220	288	378	234	322

4.3 Pre- and Post-tests

In order to analyse the games' impact on students' foreign language acquisition and communicative competence we designed a three part pre- and post-test. All students' participants were asked to fill in the tests once before and once after playing the game. Each test consists of three different exercises which all aim to test students language skills before and after playing the online-game. Whilst tests 1 and 2 focus on checking students' grammar competence (in test 1 students have to write the name of different objects and in test 2 they need to complete the missing prepositions), test 3 focuses on students' writing skills and communicative competence (students have to describe a picture and the position of different objects). Fig. 3 shows the normalized results obtained by 12 students who took three pre- and post-tests, consisting of 51 overall question items focused on, respectively, the assessment of lexical (34 items), grammatical (7 items) and written communication skills (10 items). Grades have been normalized in a 1.0-base scale. The box-plots depict how grades are improved and dispersion is reduced in all cases of assessment. A deeper statistical analysis on a greater data set is, however, required to find reasonable evidences in favor of the virtual world learning experience.

Fig. 4 shows the pre- and post-test results of grammatical and writing skills assessments with regard to those categories whose members had previously shown a low willingness to interact (i.e. category 2), and the mixed category formed by students who were and are still not willing to interact (i.e. category 3). Fig. 5 shows a comparison of overall assessment results for all categories. However, the small size of the categories and great dispersion do not allow to claim yet if grammatical or writing skills as well as students communicative competence have been improved. A higher sample size is therefore necessary to generalize the obtained results.

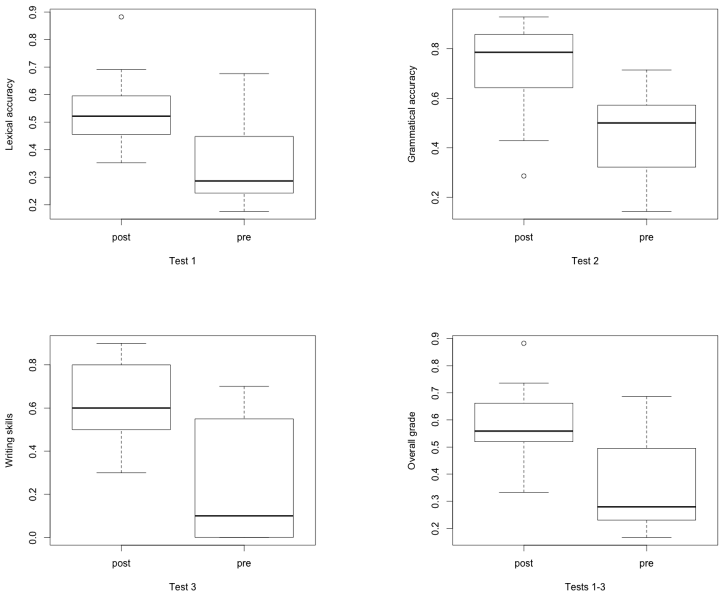


Fig. 3. Box-plots of results from the pre-test and post-test assessments

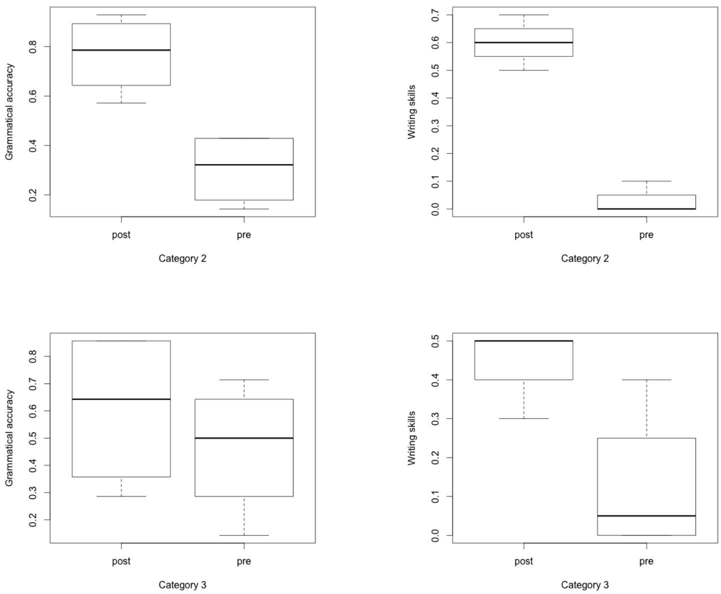


Fig. 4. Box-plots of results from the pre-test and post-test assessments for categories 2 and 3 (low willingness to interact and the mixed category)

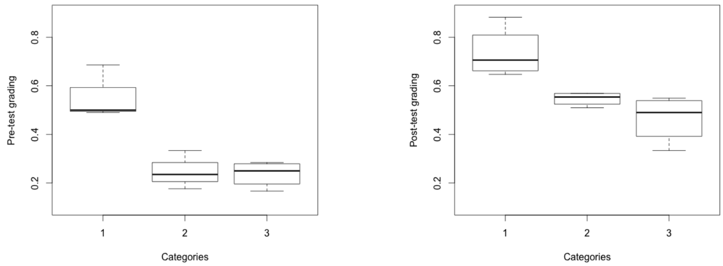


Fig. 5. Box-plots of results from the pre-test and post-test assessments of the overall grades of every category

5 Results

At first sight, we can confirm that the study has shown positive results. All students, no matter what category they belong to, increased their grades in the three aspects assessed when comparing the post-tests results with those of the pre-tests. However, results are specially interesting in terms of grammatical accuracy and writing skills and their improvement on part of some participants in categories 2 and 3 (see Fig. 4). The fact that students with a rather low willingness to interact and communicate in the target language improved significantly their language skills throughout the game, is promising. It seems that the virtual world created a new context where they were encouraged to communicate.Looking in detail at the different aspects measured in the tests, all students in category 1 obtained better results in the pre- as well as post-test than those students which were placed in categories 2 and 3 (see Fig. 5).

Regarding grammatical accuracy and writing skills, its apparent improvement does not hold for every individual, but it does for the average. It seems that there is a relationship between students' willingness to interact and communicate in the classroom and their foreign language proficiency. The number of turns, phrases, single words and total words shows a significant distortion in category 1. The first category has the highest number of all those aspects in the chat, whilst the second one has the lowest. Perhaps it highlights that the willingness to interact does not relate to accuracy in the communication. Another interesting aspect refers to the lexical mistakes with regard to students' spelling. The number of lexical mistakes in the chats was lower in categories 1 and 3. That is, when there was at least one student in the category who had previously shown a high willingness to interact and communicate in the classroom. It seems that when one student of the team takes care of the writing her mate also tries her best, resulting in low mistakes. The same applies for the use of native words and reinforcing the hypothesis (reinforcing request and confirmation checks). Another finding is that feedback is higher in categories 2 and 3. It seems that when at least one of the students in the team has not shown willingness to interact and communicate in the classroom there is more need for providing him/her with positive or negative feedback on his/her task performance.

6 Discussion and Conclusions

In the last years there has been an increasing attempt to explore the potential of computer games and virtual game-based learning environments in order to engage students towards foreign language learning. However, literature review has shown that there are still few attempts to provide empiric evidence of its educational potential, whereas our present study aims to address this lack.

In this paper we explore the possibilities of using a highly interactive 3D online game in order to reinforce students' foreign language acquisition and communicative competence at the A1 level. The game was designed by using the OpenSim platform and according to the target students' needs. In the game (*The hidden room*) students had to cooperate in teams of two in order to achieve a common goal using text chats, that were recorded and analyzed. We separated students in our pilot study in three categories, according to the willingness to interact and communicate in the target language they had previously shown within the classroom. The results obtained are promising. All students, no matter what category they belong to, increased their grades in the three aspects assessed when comparing post-tests results with those of the pre-tests. The number of lexical mistakes in the chats was lower in those categories where at least one of its members had previously shown a high willingness to interact and communicate. The same applies to the use of native words and confirmation requests. It seems that when one student of the team takes care of the writing her mate also tries her best, resulting in low mistakes. Conversely, students showing scarce willingness to interact and communicate in the classroom showed a high increase in their grammatical accuracy and writing skills. It seems that the virtual game-environment created a new context where they were encouraged to communicate.

Anyway, a deeper statistical analysis on a greater data set is necessary to find reasonable evidences in favor of the virtual world learning experience. The experimental environment can be prepared to automate text chat interaction analysis using natural language processing tools. For instance, POSIX regular expressions can help to detect regular communication patterns in text chats before being supervised to confirm or revoke its validity as evidences of communicative competence. The improvement of the research methodology based on the automated analysis of text chat data sets is an ongoing work that can provide an interesting line of further research.

Acknowledgements. This work has been funded by the Andalusian Consejeria de Innovacion Docente and the University of Cadiz under the following projects: ASCETA (Ref. P09-TIC-5230) R&D project from the Andalusian Government and (PI-13-011): Fortalecimiento de la comprension escrita en lengua extranjera mediante el uso de videojuegos: Aleman II thorugh University of Cadiz programme for Researching and Innovation in Education, and OpenDiscoverySpace:

A socially-powered and multilingual open learning infrastructure to boost the adoption of eLearning resources (European Union CIP-ICT-PSP-2011-5). We would like to thank also the Free/Libre Software and Open Knowledge Office at UCA (OSLUCA), and the game developer team Raul Gomez Sanchez (Escuela Superior de Ingenieria, UCA) and Francisco Rodriguez (Escuela Politenica Superior, UAM) as well as David Camacho (Escuela Politecnica Superior, UAM), the coordinator of the ABANT project (TIN2010-19872/TSI) and UbiCamp-526843-LLP1-2012-ES-ERASMUS-ESMO for their much valued support during the entire project.

References

1. Berns, A., Gonzalez-Pardo, A., Camacho, D.: Game-like language learning in 3-D virtual environments. Computers & Education 60(1), 210–220 (2013), doi:10.1016/j.compedu.2012.07.001
2. Berns, A., Gonzalez-Pardo, A., Camacho, D.: Implementing the Use of Virtual Worlds in the Teaching of Foreign Languages (level A1). Learning a language in virtual worlds. A Review of Innovation and ICT in language teaching methodology, Warsaw Academy of Computer Science, Warsow, 33–40 (2012a)
3. Berns, A., Palomo Duarte, M., Camacho, D.: Designing interactive and collaborative learning tasks in a 3-D virtual environment. In: Call: Using, Learning, Knowing, EUROCALL Conference Proceedings, pp. 20–25 (2012b)
4. Broma, C., Preuss, M., Klement, D.: Are educational computer micro-games engaging and effective for knowledge acquisition at high-schools? A quasi-experimental study. Computers & Education 57, 1971–1988 (2011)
5. Bytheway, J.: Vocabulary learning strategies in massively multiplayer online role-playing games: A thesis submitted to the Victoria University of Wellington in partial fulfilment of the requirments for the degree (unpublished Masters Thesis) (2011)
6. Connolly, T.M., Stansfield, M., Hainey, T.: An alternate reality game for language learning: ARGuing for multilingual motivation. Computers & Education 57, 1389–1415 (2011)
7. Cornillie, F., Thorne, S., Desmet, P.: Digital games for language learning: from hype to insight? ReCall 24(3), 243–256 (2012)
8. Del Blanco, A., Eugenio, J., Marchiori, E., Fernández-Manjón, B.: Adventure Games and Language Learning. In: First International Workshop on Technological Innovation for Specialized Linguistic Domains: Theoretical and Methodological Perspectives (TISLID 2010), Madrid, Spain, October 21-22, pp. 243–254 (2010)
9. De Haan, J., Reed, W.M., Kuwada, K.: The effect of interactivity with a music video game on second language vocabulary recall. Language Learning & Technology 14(2), 74–94 (2010)
10. Di Blas, N., Bucciero, A., Mainetti, L., Paolini, P.: MultiUser Virtual Environments for Learning: Experience and Technology Design. IEEE Transactions on Learning Technologies 5(4), 349–365 (2012)
11. Doughty, C., Pica, T.: Information gap tasks: Do they facilitate second language acquisition? TESOL Quarterly 20(2), 305–325 (1986)
12. Garcia-Montero, A., Rising, B., Watts, F.: Simulation Gaming and the acquisition of communicative competence in another language. Simulation & Gaming 32(4), 481–491 (2001)

13. Gass, S.M., Varonis, E.M.: Conversation interactions and the development of L2 grammar. Studies in Second Language Acquisition 16 (1994)
14. Krashen, S.: The input hypothesis. Longman, New York (1985)
15. Long, M.H.: The role of the linguistic environment in second language acquisition. In: Ritchie, W.C., Bhatia, T.K. (eds.) Handbook O/Research on Language Acquisition. Second Language Acquisition, vol. 2, pp. 413–468. Academic, New York (1996)
16. Mawer, K., Stanley, G.: Digital Play. Computer games and language aims. Peaslake Delta Publishing (2011)
17. Marty, J.-C., Carron, T.: Observation of Collaborative Activities in a Game-Based Learning Platform. IEEE Transactions on Learning Technologies 4(1), 98–110 (2013)
18. Oxford, R.L.: Cooperative Learning, Collaborative Learning, and Interaction: Three Communicative Strands in the Language Classroom. The Modern Language Journal 81(4), 443–456 (2004), Special Issue: Interaction, collaboration, and Cooperation: Learning Language Teachers (Winter 1997)
19. Peterson, M.: Computerized Games and Simulations in Computer-Assisted Language Learning: A Meta-Analysis of Research. Simulation & Gaming 41(1), 72–93 (2010)
20. Pica, T.: Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? Language Learning 44(4), 493–527 (1994)
21. Rama, P.S., Black, R.W., Van Es, E., Warschauer, M.: Affordances for second language learning in World of Warcraft. RecALL 24(3), 322–338 (2012)
22. Reinders, H., Wattana, S.: Learn English or die: The effects of digital games on interaction and willingness to communicate in a foreign language. Digital Culture & Education 3(1), 3–29 (2011)
23. Reinders, H. (ed.): Digital Games in Language Learning and Teaching. Palgrave Macmillan, Basingstoke (2012)
24. Schmidt, R.W.: The role of consciousness in second language learning. Applied Linguistics 11(2), 129–158 (1990)
25. Sylven, L., Sundqvist, P.: English L2 learning and L2 proficiency among young learners. ReCALL 24(3), 302–321 (2012)
26. Sykes, J.M., Reinhardt, J., Thorne, S.L.: Multiuser digital games as sites for research and practice. In: Hult, F.M. (ed.) Directions and Prospects for Educational Linguistics, pp. 117–135. Springer, Amsterdam (2010)
27. Schwienhorst, K.: Learning a second language in three dimensions: Potential benefits and the evidence so far. Themes in Science and Technology Education 2(1-2), 153–164 (2009)
28. Stenlev, J.: Cooperative Learning in foreign language teaching. Sprogforum Nummer 25, 33–42 (2003)
29. Swain, M.: The output hypothesis: Theory and research. In: Hinkel, E. (ed.) Handbook on Research in Second Language Teaching and Learning, pp. 471–484. Lawrence Erlbaum, Mahwah (2005)
30. Thorne, S.L., Black, W.R., Sykes, M.S.: Second Language use, socialization, and learning in Internet Interest Communities and Online Gaming. The Modern Language Journal 93, 802–823 (2009)
31. Warschauer, M.: Computer-mediated collaborative learning. Theory and Practice. The Modern Language Journal 81(4), 470–481 (2011)
32. Zuo, W.: The Effects of Cooperative Learning on Improving College Students' Reading Comprehension. Theory and Practice in Language Studies 1(8), 986–989 (2011)

Analysis of Learners' Fieldtrip Talk during a Collaborative Inquiry Task

Canan Blake, Eileen Scanlon, Alison Twiner, Trevor Collins, Ann Jones, and Lucinda Kerawalla

The Open University, Walton Hall, Milton Keynes, UK
{canan.blake,eileen.scanlon,t.d.collins,ann.jones, cindy.kerawalla}@open.ac.uk, Alison@twiner.net

Abstract. In this paper we analyse children's talk with a view to understand how a technology enhanced inquiry learning toolkit played a part in enriching collaboration during a fieldtrip and facilitating social interaction. The participants in the study were 15 year-old students carrying out their geography GCSE (General Certificate in Secondary Education) work in a secondary school in the UK. During the fieldtrip, we provided students with nQuire, an inquiry learning toolkit to orchestrate their learning, on an ultra-mobile Asus Eee PC with a wireless connection to the coursework web site. Students collected data from twelve points in two towns with very different layout and land use. The learning environment created with the nQuire toolkit, ultra mobile PCs, personalised inquiry task and the use of scientific sensors to collect data offers possibilities for collaboration and effective interaction. In this study we analyse to what extent this environment supported learning collaboratively and to what extent students interacted with each other and with the technology to construct knowledge during the fieldtrip.

Keywords: collaborative learning, CSCL, inquiry learning, nQuire.

1 Introduction

Computer supported collaborative learning (CSCL) environments are based on the premise that learners will work together as a group to achieve the given task and engage in productive interactions that may lead to learning. For example, Dillenbourg et al (1) argue that in order for collaboration to produce learning outcomes the groups should actually engage in productive interactions. They propose three main categories of interactions that have been found to facilitate learning: explanation, argumentation/negotiation and mutual regulation. They further argue that "the purpose of a CSCL environment is not simply to enable collaboration across distance but to create conditions in which effective group interactions are expected to occur" (p. 6).

Given that the main focus of CSCL research is "on how people learn in the context of collaborative activity and how technological settings that support this activity can be designed and evaluated" (2) as well as looking at interactions among learners that gives rise to collaborative learning (3), we need to pay attention to technological tools

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 32–42, 2013.

we provide in these environments as they will mediate and shape the nature of the interaction among learners.

When students engage in discussions regarding the subject they are studying, their turn-by-turn exchanges can provide us with information about the learning taking place. These discussions occur when students explain to each other their understanding, ideas and strategies, ask questions, defend their perspective and evaluate each other's suggestions. In order to increase the likelihood of productive interactions among learners, researchers, especially those working in the CSCL field, have been implementing particular applications of technology to benefit learners and facilitate teaching. Lipponen (4) emphasizes that collaborative learning supported by technology can enhance peer interaction and work in groups. The guiding framework for implementation of technology in CSCL environments is socially oriented theories of cognition and learning.

Guided by these approaches, in this paper we will present an analysis of children's fieldtrip talk during a geography inquiry task. We used a detailed video analysis approach based on the work of Ash (5) and our aim is to understand how students made use of the tools, technologies and inquiry framework during this task.

We will first present the context of the study, including the Personal Inquiry project and the field trip. Then we will describe how we analyzed the video recordings of selected groups to help us understand how students collaborated to achieve the task and how technology mediated the social interactions during the field trip.

2 Research Context and Data Collection

2.1 Personal Inquiry Project and Urban Heat Islands Field Trip

The aim of the personal inquiry project was "to support children aged 11-15 years in coming to understand themselves and their world through scripted personal inquiry learning" (6). Children investigated issues that affect their lives, across different settings - including the classroom, their homes, and nature centres - through a scientific process of gathering and assessing evidence, conducting experiments and engaging in informed debate. A computer toolkit, named nQuire, was designed to enable scripted inquiry learning, where scripts are computer programs, like dynamic lesson plans, that guide and support the learners through an inquiry learning process by providing them with a set of structured activities, data probes, visualizations of data, and means of communication.

In one theme, explored in the school based trials, 135 students investigated Urban Heat Islands at Milton Keynes and Northampton. In this paper we will report on the analyses of the fieldtrip video recordings with an aim of understanding the nature and quality of collaborative actions.

The participants in the study were 15 year-old students carrying out their geography GCSE (General Certificate in Secondary Education) work in a secondary school in the UK. The fieldtrip presented in this paper was part of the course-work towards this qualification. In the fieldtrip, students collected data across two towns (Milton Keynes and Northampton) to investigate the urban heat island phenomenon. The towns are different from each other in terms of their age and layout of the buildings.

We provided students with nQuire inquiry learning toolkit to orchestrate their learning (supporting hypothesis generation, experimental design, data collection analysis, presentation). See screen shot in Figure 1.

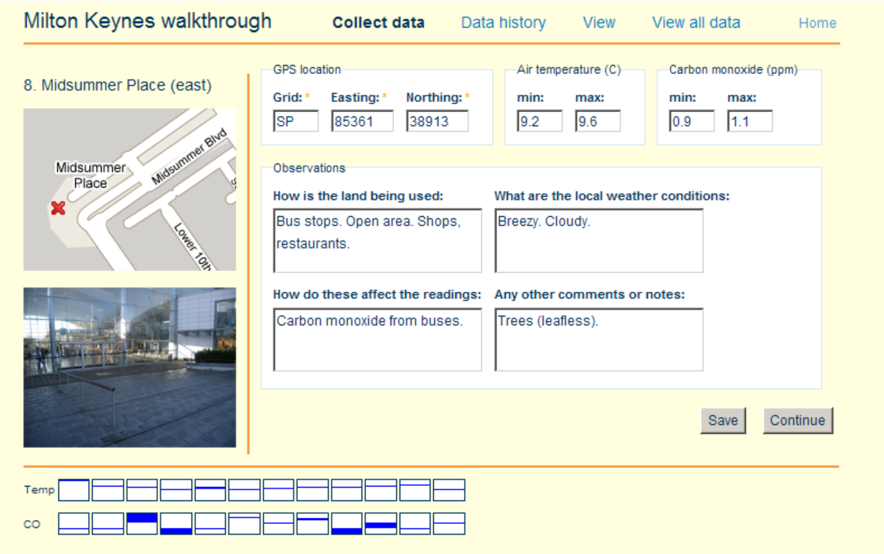


Fig. 1. Screen shot of data entry screen in nQuire

Students collected data using sensors provided by our industry partners (ScienceScope: http://www.sciencescope.co.uk/index.html) attached to their mini laptops. They collected data from 12 points in each town during the day and entered their data (air temperature, GPS location, photographs of the data collection spot and either Infra Red (IR) radiance, wind speed or carbon monoxide levels) into nQuire. They were using Asus Eee netbooks to access nQuire to support their fieldwork data collection. nQuire data entry screen structured the students' data collection activities and at the same time helped students to deal with real data by stipulating which types of data was acceptable for each variable.

We video recorded two focal groups in the first year and four focal groups in the second year of the study (4-5 students per group) during their field trip. Students' talk was transcribed and analysed using video analysis approach used by Ash (5) and then a qualitative content analysis was employed to code utterances. The unit of analysis was defined as the single contribution made by a student at one time.

3 Analysis

In this study we employed the video analysis approach used by Ash (5). She studied the informal science learning in settings such as museums, after-school programs and community technology centres. She used video tools for analyzing discontinuous dialogue and used the three levels of analysis which provided the means to analyze complex sets of dialogic data without oversimplifying their meaning (p. 225). With this three level analysis she managed to balance complexity and reductionism to get a more accurate and productive data analysis (p210). The three levels of data analysis consist of:

1. The flow chart that gives an overview of one entire visit, as well as the pre & post interviews.
2. The significant event consisting of one segment of the flow chart with productive exchanges and analysed in greater detail.
3. The microgenetic analysis comprising a detailed dialogic analysis of the SE in focus.

The first level of our analysis, what Ash (5) called 'a flow chart' was an overview of the entire fieldtrip carried out by the group. We identified the segments of the activities that can be analysed in more detail later and presented them in the order in which they happened. A typical section from the fieldtrip flow chart is presented in Figure 2 indicating the location of data collection point and associated notes for this location. The first column indicates the lines in the video transcript.

UHI – study 1 Group CPDFJ			
Lines	Location	Overview	Content themes
1-90	NH-1 Station	Data collected by helping each other, everyone contributed, making notes of taxis, and temperature in different parts…	• Figuring out which is easting and northing for GPS • Finding out what order to do readings in • Photos for social purpose on trip as well
91-134	NH-2 Main road to town	Collecting data at a bus stop, noticing traffic light at a nearby junction, talking about pollution…	• Pollution • The effect of traffic light on air pollution • Finding the best position for measuring wind speed

Fig. 2. Part of flow chart for Group CPDFJ

The next level of analysis, called significant event (SE) by Ash (5) is a more detailed analysis of a segment identified in the earlier analysis paying attention to dialogue, content of the activity, and the tools and technologies employed during the interaction to explore the students' understandings of the scientific content and to make connections to their prior knowledge.

According to Ash each SE is just large enough to encompass ‘one discussion’ and contains:

- o recognisable beginnings and endings, generally but not always centred on one topic;
- o sustained conversational segments that differ from the short, un-sustained interactions that can precede and follow SEs;
- o different sources of knowledge, such as distributed expertise; and
- o inquiry strategies such as questioning, inferring or predicting (p. 216).

We present a SE segment below which is representative of student interactions, illustrating how learners make sense of their observations and negotiate their understanding with others in their group, and how they negotiate tasks. The number of SEs identified in the groups ranged from ten to four.

Group CPDFJ	SE4		
Lines	Location	Topic	Detailed context
148-171	NH-3	Land use, weather conditions and tall buildings	Students notice they are on a main road with a church and a Library nearby. They all talk in a complementary way by adding their observations to previous students’ observation. The weather is sunny, windy and very cold. One of them suggests the wind is tunnelled through the street. Another student takes a picture and shows it to the others. They also decide it is fairly open and discuss if it makes sense. One of the quieter students observes that the buildings get bigger as they get nearer the centre of town.

Fig. 3. An example of a Significant Event (SE)

The next stage of analysis is the fine grained analysis of interaction in these selected segments. At this point we also added our own approach to fine grained analysis of collaborative talk (see 7). We employed a classification system developed in our earlier work (8,7) and validated and extended it using the rating scheme by Meier et al. (9) for assessing the quality of computer supported collaboration processes (see 9,10). For this purpose we marked all student talk in these SEs into one of the following categories:

1. Joint knowledge building utterance - JKB
2. Motivation and commitment utterance - M
3. Task organization/information utterance - TO
4. Data collection utterance - D

Joint knowledge building segments contains dialogue where students present and defend a position, give an example, ask questions and explain.

For example:

C: What does infra-red irradiance actually mean?

D: It's the amount of light reflected… So, have we finished then? Has Mary taken a photo?

Another example:

C: [to H] Are you going to take it off the tree?

D: Might as well, there's no other material around here that we haven't already done, is there?

R: But it's not the material, it's the local conditions, isn't it?

D: Yeah, we're just testing, doing a comparison between the materials as well.

Task organization/ information segments are contributions that are related to co-ordination of the collaborative activity.

J: I've worked out, I've worked out that for 24, so everyone's gotta swap at 6. So then everyone can have a go at every tool

Y: OK

J: So at 6 I'll have that (points at camera)

V: Yeah

J: And then you can have GPS or the laptop. And then P has the laptop and Y has this thing (datalogger).

Motivation and commitment segments are about participants' individual commitment and motivation.

J: I can't spell today

V: Oh hurry up I'm holding my hand up here

J: Right I've done it, done it. Wait, wait, wait, wait. I've gotta go to, eastings.

Another example:

V: Oh I'd love to live up there
F: It'd probably be a bit noisy
V: Yeah but I'd

Data segments contain data collected to accomplish the task. (This category is specific to the analysis of field trip data presented in this paper; it is not treated as a collaborative category although they have to agree how to carry out the activity as a group as there are several data loggers and other equipment like the laptop and the camera).

D: Minus 28 (laughs)
L: Bit cold innit?
D: Miss it says minus 28. Oh it's gone up to 13 now.
L: (reads) 13.6, 14, hang on a minute.
D: 14, and 13.8
J: I got an Anglia water van (picture)
L: (talks when typing)
J: Loads of traffic.

The fine grained analysis of video transcript for the above SE is presented below. The SEs mostly contained joint knowledge building type contributions as these were the segments identified as showing children exchanging, explaining ideas related to the features of two towns. After categorizing children's talk in SEs we applied the same analysis to the rest of the transcript for all groups in order to get an idea of type of interactive talk students had on the field trip.

Group CPDFJ	**SE4**	
Lines		**Codes**
148-171	Peter: OK. How is the land being used? (**Q-JKB**) Carol: Main road and (**A-JKB)** Felicity: Is that a library? (**Q-JKB**) Carol: But there's a church opposite as well. (**A-JKB**) Joel: Road (**A-JKB)** Peter: What's the local weather conditions? (**Q-JKB**) Carol: still sunny. Not many clouds. (**D**) Felicity: no clouds (**D**)	Question – JKB: **Q-JKB** Answer – JKB: **A-JKB** Data-JKB: **D-JKB** Data only: **D** Task organization :**TO** Motivation: **M**

	Peter: (typing) 'no clouds'. There's quite a bit of wind isn't there. (**D-JKB)** Felicity: Yeah (**D-JKB)** Peter: (typing) 'it's still cold' (**M)** Carol: It's kind of tunnelled through the building entrance (**JKB)** Peter: Yeah. How does this affect the readings? (**Q-JKB)** Carol: Erm. Cold and the wind's being tunnelled through the street (**A-JKB)** Peter: OK so tunneled (**JKB)** Felicity: I've got a picture (shows to Candice) (**D)** Carol: Oh wow (**M)** Peter: Tunnelled through street. Any other comments or notes? **(JKB)** Carol: Erm, it's freezing **(JKB)** Peter: It's freezing **(JKB)** Felicity: It's fairly open **(JKB)** Peter: Yeah that doesn't make sense **(JKB)** Felicity: But it is. **(JKB)** David: The buildings are getting bigger as we get nearer the centre of town. **(JKB)**	

Fig. 4. Fine grained analysis of the dialogue in the SE presented in Fig. 3

4 Discussion

This three level analysis of children's talk during a geography field trip enabled us to identify points at which learners worked together and the kind of interactive exchanges took place, i.e. productive exchanges to achieve the given task or just talk about non-task subjects. We were also able to identify the role of technology, in this case nQuire toolkit, in supporting these exchanges among learners.

The first level of analysis provided us with an overview of the children's field trip. The second level analysis, identifying Significant Events (SEs), enabled us to recognize important episodes of collaboration among children, how they worked out the task organization, the contribution of the task set-up in getting mostly everyone to contribute to the task, and also the effect of (individual) motivation on the task.

The third level analysis provided us with a fine grained presentation of the children's collaboration. In addition, applying this fine grained analysis to the whole of the children's talk during the fieldtrip enabled us to present a different overview of the fieldtrip as shown in Figure 5. The chart presents the percentage of their talk the categories discussed earlier for each group. As can be seen from the chart, most of the students' exchanges are in the data category. These are short exchanges between students associated with the values read from the data sensors they use to determine temperature, GPS location, wind speed, IR radiance or CO level. Although they are numerous and shown as the most frequent category, given that they are short exchanges, it is possible to argue that in terms of time spent, this data category might not be the most time consuming.

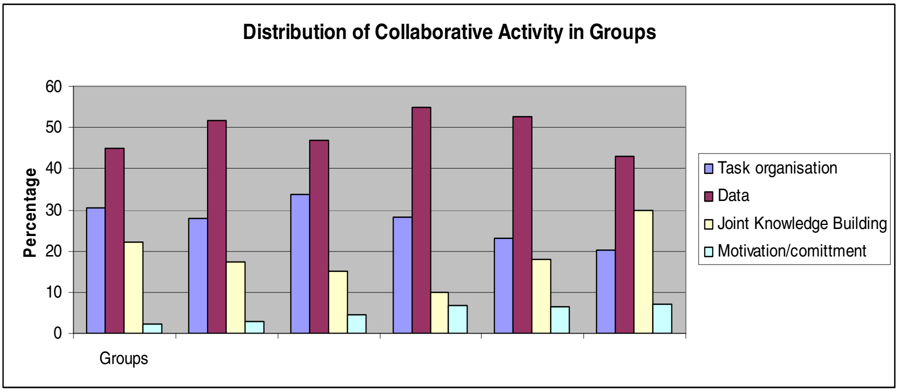


Fig. 5. Distribution of Collaborative Activity in Groups

When we look at the **Task organization** category we can see that the groups are different from each other; some groups spent about 35% of their total talk in organising the activity, whereas others spent about 20% of their total talk. **Task organization** was quite central to some groups as they were discussing who would be responsible for what type of data, when the data sensors, mini laptop and camera would be swapped between students and what was the best way of doing this. Some other groups decided who would collect which data at the beginning of the day and carried out with the same students until the end of the activity. They also supported each other as entering the data into the laptop while standing was not easy, especially in the cold weather.

The **Joint knowledge building** category contains students' collaborative talk related to the concepts they come across while collection data and observing land use and traffic in the area. These are exchanges that show that the field trip data collection had collaborative episodes where children asked each other questions about the phenomena they observed, explained these using their observation and prior knowledge and answered questions asked by the other group members.

5 Conclusion

Children in an exploratory fieldtrip have to deal with multiple activities, and may not collaborate all the time. So tasks need to be designed in a way to allow both individual and collaborative activity and also they need to be designed with flexibility so that learners will be able to perform both individual and collaborative tasks. In this study we were able to observe how the design of the task enabled students work together, solve problems together, explain each other unclear points and decide how to achieve the given task. The nQuire data entry screen structured the students data collection activities, provided a point of focus for the interaction among students and at the same time helped students to deal with real data by stipulating which types of data were acceptable for each variable.

Task organization worked differently for everyone and each group found their own way of agreeing what to do. Consequently there were some groups having lengthy discussions about how to do it as seen in the student exchange under Task organization/information category). In all groups students worked together and shared the task.

It was observed that some groups worked in a harmonious way and efficiently. In some groups students said thank you to each other after given data/information. They supported each other during the fieldtrip and worked collectively to achieve the task as can be seen in the following exchanges:

Em: No, no, no, it's OK. (typing) 'buildings emit heat and

Li: Do you want me to hold it while you like type with both hands? Might be a bit easier for you.

Lu: How long have we got?

Em: Two more stops.

...

S: The glass building is made out of dark coloured bricks and glass

A: Yeah but that's for the infraradiance

N: Also the chimneys over there (points)

A: OK (starts typing)

S: Also the cars

A: Yeah

S: It's not rush hour, so

A: Did you take a picture of that smoke? (points to chimneys)

(S takes picture of chimney, A types)

S: And then put it's like near industries that let off smoke

B: Yeah, which would make the carbon monoxide (inaudible, 04.15).

These groups had more SEs identified in their transcripts. We observed that having an individual and/or collaborative commitment to task increases student interaction and they complete the task more efficiently. However for some groups the cold weather and the long day outside were too much and they did not do much apart from collection the data as quickly as they could. Some data sensors were not popular with everyone and these groups negotiated, in length, changeover points for these instruments.

The three-level approach to analysis of video transcripts provided us with both an overview and a detailed account of important points with respect to learners' interactions. As a midway point Significant Event (SE) segments provided representative sections from the interaction of students which provides information about how students negotiate the scientific concepts, task organisation and data collection in the field trip.

Our future plans are to explore the possibility of integrating the analysis methodology (or part of it as this may be the case) to online environments such as nQuire so that student interactions can be analysed more easily. This can be achieved for example by prompting students during the activity to make notes of the important events every half hour or so. This would also help teachers to monitor the class activity in regular intervals and would help support the collaborative work of small groups working simultaneously, especially in online environments. This may enable researchers and teachers to track collaboration more accurately and in real time.

Acknowledgements. We would like to acknowledge the contributions of students and teachers in the secondary school in Milton Keynes and all members of the PI project.

References

1. Dillenbourg, P., Jarvala, S., Fischer, F.: The evolution of research on Computer Supported Collaborative Learning. From design to orchestration. In: Balacheff, N., et al. (eds.) Technology Enhanced Learning. Principles and Products, pp. 3–19. Springer (2009)
2. Overdijk, M., van Diggelen, W., Kirschner, P.A., Baker, M.: On the connection of agents and artifacts in CSCL: Toward a rationale of mutual shaping. International Journal of Computer Supported Collaborative Learning 7(2), 193–210 (2012)
3. Dillenbourg, P.: What do you mean by collaborative learning? In: Dillenbourg, P. (ed.) Collaborative-learning: Cognitive and Computational Approaches. Pergamon/Elsevier Science (1999)
4. Lipponen, L.: Exploring foundations for computer supported collaborative learning. In: Stahl, G. (ed.) Computer Support for Collaborative Learning: Foundations for a CSCL Community, pp. 72–81. Lawrence Erlbaum, Hillsdale (2002)
5. Ash, D.: Using video data to capture discontinuous science meaning making in non-school settings. In: Ed Video Research in the Learning Sciences. Peter Lang Press (2007)
6. The Final Report of Personal Inquiry (PI): Designing for Evidence-based Inquiry Learning across Formal and Informal Settings, `http://www.pi-project.ac.uk/news/`
7. Blake, C., Scanlon, E.: Analysing Collaborative Processes and Interaction Patterns in Online Discussions. In: Eighth International Conference on Networked Learning, April 02-04. Maastricht School of Management, Maastricht (2012)
8. Jones, A., Scanlon, E., Blake, C.: Conferencing in communities of learners: Examples from social history and science communication. Educational Technology and Society 3, 215–226 (2000)
9. Meier, A., Spada, H., Rummel, N.: A rating scheme for assessing the quality of computer-supported collaboration processes. International Journal of Computer-Supported Collaborative Learning 2(1), 63–86 (2007)
10. Rummel, N., Deiglemayr, A., Spada, H., Kahrimanis, G., Avouris, N.: Analysing Collaborative Interactions Across Domains and Settings: An Adaptable Rating Scheme. In: Puntambekar, S., Erkens, G., Hmelo-Silver, C. (eds.) Analyzing Collaborative Interactions in CSCL: Methods, Approaches, and Issues. Springer, New York (2011)

Teaching Low-Functioning Autistic Children: ABCD SW

Maria Claudia Buzzi, Marina Buzzi, Beatrice Rapisarda,
Caterina Senette, and Maurizio Tesconi

CNR-IIT, via Moruzzi 1, 56124 Pisa, Italy
{Claudia.Buzzi,Marina.Buzzi,Beatrice.Rapisarda,
Caterina.Senette,Maurizio.Tesconi}@iit.cnr.it

Abstract. Applied Behavior Analysis (ABA) is highly effective for teaching subjects with autism. We discuss the design and development of a free open-source customizable software to support ABA intervention in low-functioning children with autism. The software automates the trial setting while enabling the gathering of the children's performance data to monitor learning. The software relies on a Web architecture: using a laptop, the tutor defines the exercises dynamically activated on the child's tablet. Synchronization between these two devices occurs via an Internet connection, obtaining and inserting data through the database. A real-time summary of the actions performed by the child is available on the tutor's device, simplifying decisions about the intervention. In order to make the trial accessible to any child, the software adapts the visual prompt to the child's abilities, i.e., receptive/verbal, using labels in addition to images. A pilot test with several children confirmed the usability of the software.

Keywords: Didactic software, Web system, Autism, Applied Behavior Analysis, Augmentative and Alternative Communication, Discrete Trial Training.

1 Introduction

Autism Spectrum Disorder (ASD) is a pervasive developmental disorder characterized by difficulties in communication and social integration, and by the presence of behavioral problems, although symptoms vary significantly from child to child. Since traditional educational methods are rarely effective in ASD, new teaching approaches aimed at better exploiting the subject's abilities are currently adopted. Of these, Applied Behavioral Analysis (ABA) is a scientific approach with proven effectiveness for teaching low-functioning children with autism [8,9,12]. It provides one-to-one teaching via trials of increasing difficulty according to Discrete Trial Training (DTT) and Augmentative and Alternative Communication (AAC) [5] in errorless learning. The child's progress is constantly monitored by collecting data from the child's performance.

In this paper we describe ABCD SW (Applied Behavior and Computer-based Didactic SoftWare), a free open source software to support learning in low-functioning autistic children through ABA intervention. Even learning about simple articles such as common everyday objects or relatives can be very difficult for low-functioning children. The didactic software implements basic modules to teach common information -- articles,

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 43–56, 2013.

foods, animals, numbers, letters, places, relatives, human body parts, etc. – in order to stimulate receptive and expressive language. ABCD software is customizable; for instance, non-verbal children are supported by the written word and it is possible to define the most appropriate multimedia reinforcements.

The software relies on a Web architecture: using a laptop, the tutor defines the exercises dynamically activated on the child's screen. The child interacts with the software through a touchscreen tablet since tactile interaction is more natural and easily used by 3- to 6-year-old children; however, the child could use any touchscreen device provided with an Internet browser able to support Javascript and HTML5. The software automates the trial setting while it enables gathering the children's performance data to monitor the learning. Section 2 introduces related work, while Section 3 shows the architecture of the system and the user interfaces, discussing design choices as well as problems encountered and solutions adopted. Afterwards, Section 4 describes the pilot test, focusing on the software improvements and lessons learnt. Section 5 shows results from a qualitative assessment of the software. Conclusions and future work end the paper.

2 Background

2.1 Related Work

Research on the potential of CAI (Computer Assisted Instruction) in autism began in the 1970s, and results show that associated learning scenarios are easier for people with autism thanks to the emotional distance between a PC and the user, which can reduce anxiety and problematic behaviors [11]. In recent years, many programs for autism have been developed, but only a few are specifically designed to support ABA methodology. A team from NUI Galway has created a web engine that could be helpful for developers of applications based on ABA principles, but unfortunately this proposal relies on proprietary technology with considerable development costs [6]. Furthermore, nearly all the software currently available is commercial, and thus not economically feasible for all families of children with autism. A 2006 study by Whalen et al. based on the use of software designed on ABA principles showed that the use of CAI improves social and communication skills by reducing inappropriate behavior [13]. Another study showed that computer instruction is more effective when combined with the physical presence of the tutor rather than when used alone [10]. In agreement with these results, our study favors a platform where tutor and child work together while using two separate areas to interact with the same application.

Since a chance to learn is directly linked to the possibility of communicating, in recent years much effort has been made to develop tools that facilitate communication by means of technologies based on Augmentative and Alternative Communication (AAC). Recent research has confirmed the benefits of electronic intervention based on AAC. Hirano et al. [4] designed and implemented a visual scheduling system (vSked) for planning and organizing children's activities in the classroom in a visual way. In addition to increased efficiency for caregivers, they also observed improvements in student-student and student-teacher communication. Monibi and Hayes have conducted comparative studies on current existing applications in order to classify the characteristics in terms of requirements and constraints, proposing different approaches for future communication tools. Their prototype tool, MOTOCOS, has been proposed to overcome the limitations of existing tools [7].

According to teachers and experts, the available digital tools (e.g., Tango[1], Activity Pad[2], DynaVox[3]) still require expertise and training, showing little usability and flexibility. Recently, a qualitative study in 2010 focused on several available technologies (vSked, Motocos, SensaCam Microsoft[4]), identifying their main limitation as the inability to document and monitor over time any use and progress in the learning pro cess. Considering these findings, the authors then redesigned these applications as needed and defined new guidelines to combine the advantages of analog devices with the potential of ubiquitous computing solutions [3]. Conversely, the ABA approach considers evaluating the subject's progress crucial to defining future learning programs.

Research in this field is currently focusing on using mobile devices. There are several interesting studies [2, 11] specifically focused on communication and language but to the author's knowledge, these applications do not also monitor the learning trends, one of the main goal of ABA. Instead, our didactic software records all relevant data for providing effective monitoring of children's learning and behavior to the educators, supporting an accurate assessment of the intervention. Furthermore, our software is open source and free, relieving families of additional costs for software (http://abcd.iit.cnr.it/). Free software has more chance of being used by many people and we can take advantage of that. Open source software can benefit from contributions from the developers' community, which could help us implement new functionalities to cover different ABA programs.

3 The ABCD System

ABCD SW is a web-based application for teaching low-functioning autistic children as briefly described in [1]. It is based on ABA intervention, which provides one-to-one errorless teaching with trials of increasing difficulty according to Discrete Trial Training (DTT) and Augmentative and Alternative Communication (AAC). A child's progress is constantly monitored by collecting data on the child's performance in each trial executed during a session, e.g., category, article, correctness of the trial, type and the percentage of prompt, comments. To better understand the design process, we present some basic concepts. ABA intervention is conducted in a simple environment that offers no distractions for the child and takes place on an empty table using cards, photos or real objects. The tutor selects the items needed to set up a specific trial (the article to be learnt, and possible additional materials). The trial's set-up takes time, also considering the time needed to prepare all the objects beforehand. The teacher, after preparation for the child's trial, announces the command of the program to be executed. The basic programs of the learning process are: 1) matching, to identify the

---

[1] http://www.tango.me/
[2] http://www.activitypad.com/
[3] http://uk.dynavoxtech.com/default.aspx
[4] http://research.microsoft.com/en-us/um/cambridge/projects/sens ecam

image, and possibly the label, representing the object; 2) receptive, to verify whether the child correctly identifies the object; 3) expressive, to solicit communication regarding the object. In ABA each program is implemented by means of a sequence of trials of increasing difficulty, according to the DTT.

ABCD SW automates three phases of the ABA intervention, facilitating and speeding up the tutor's work: 1) the trial set-up, allowing easy addition of items and categories by using a CMS, 2) configuration and control of the flow of the educational programs assigned to the child according to the ABA sequence and 3) evaluation of each trial executed.

An ABA team usually consists of many tutors, including the child's parents, to guarantee consistency in the behavioral approach. More than one tutor is necessary in order to ensure that the child's learning is not depending on a specific person.

ABCD SW was developed using the participatory design approach, to better guarantee the efficacy and usability of the tools realized. Examples of applied participatory design for software development are presented in [3] and [7]. The software design was then based on a multidisciplinary team comprising software analysts, designers and developers, two psychologists, six ABA tutors, an ABA consultant and two parents of children with autism. ABA principles have been mapped into the ABCD SW in different ways:

- Comprehensible: the software is based on Augmentative and Alternative Communication, so it uses images (also animated) and labels.
- Errorless: the software tries to avoid a child's error (see Section 3.4).
- Learning by increasing difficulty levels. This requirement has been satisfied implementing the Discrete Trial Training defined in the ABA, including mass trials, trials with one or two (neutral and not) distracters, extended trials and rotations.
- Monitoring children's progress. ABCD SW records data on the trials regarding performance and evaluation. This point is central and focal to the ABA since it allows the interventions and progress to be measurable, favoring the set-up of the teaching strategy according to results.

3.1 The Architecture

ABCD SW is a dynamic Rich Internet Application based on distributed Web architecture, with two devices for the tutor and the child -- a laptop and a touchscreen tablet -- connected via Internet. No physical connection between the devices is required, making interaction more natural and suitable for autistic children, avoiding the possibility that they may be attracted to plug in/out cables. Using a laptop, the tutor defines the program and level of the ABA exercises to run the trial, which is dynamically activated on the screen of the child's tablet.

Specifically, ABCD SW is a PHP, AJAX and HTML5 Web application that relies on the Drupal Content Management System (CMS) and the mySQL database. The use of the Content Management System enhances data management, offering native scalability and support for internationalization. Most of all, it permits managing application content independently from the code, by allowing end users to add/modify content in a user-friendly way. All the software's functions are implemented using jQuery and JSON, which allow the efficient and easy exchange of data. The AJAX technology, handling calls to the server, allows the HTML5 client user interfaces to be constantly updated. Images extracted from the database are dynamically inscribed in rectangles and carefully designed to be simple and not a source of distraction. For drawing objects in the Web browser, preferably Google Chrome or Mozilla Firefox since they better support HTML5 functionalities, the JavaScript Library Raphaël using SVG (Scalable Vector Graphics) has been adopted. The library allows easy positioning of the image on the "canvas" and provides methods for animating the element, enabling its movement when needed. The use of SVG makes it possible to adapt the canvas to different screen sizes respecting the relative distance between displayed items, avoiding any confusion for the child.

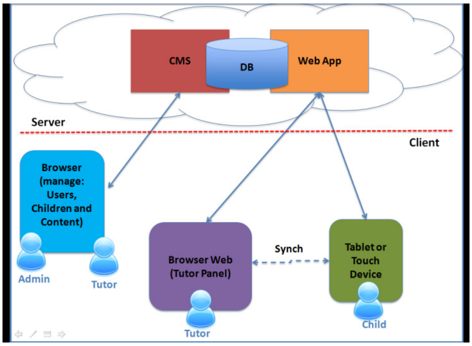


Fig. 1. The Architecture

The ABCD SW architecture considers the devices involved as autonomous and independent. This independence is guaranteed while making full use of their Internet connectivity. The communication modality between the devices is created using the database as a common channel. Each event from the tutor's or the child's device is saved on the channel, and every second the two terminals are synchronized with new data (if present). Specifically, the mechanism has been implemented with multiple jQuery's AJAX call to the server based on AJAX long poll technique. An alternative for the server-client conversation could have been the use of HTML 5 sockets but at the moment the specifications of this technology do not seem to be sufficiently stable. The effect of this implementation can be seen in the Tutor Control Panel (Fig. 4), the interface from which the tutor can see all information regarding what is happening on the child's screen.

3.2 The Tutor User Interface

Two main user interfaces (UI) are provided for the tutor: 1) administration, for adding a new child, changing the configuration and defining reinforcements; 2) working, for starting the ABA session and recording data.

To access the didactic software, the tutor has to log-in and select the child's nickname -- several children can be followed by the same tutor. The tutor can use the administration UI to add one or more children, specifying in the configuration file their ability (receptive/not, expressive/not) to adapt the software to their needs. The child will be automatically related to the tutor; it is possible to relate a child to more than one tutor. After log-in and selection of the child's nickname, a summary presents data on all the previous trials/sessions performed by the child with the software. If needed, the tutor can read information on the recent trials performed by different tutors, or start a new session by pressing a push button that activates the tutor environment (Fig. 2). If the child has already used the software, the system proposes the last trial performed, i.e., the last program-category-level-article at which the child has worked. All the articles of the category are visible as images grouped by the current learning status (mastered, not mastered, on acquisition, excluded, suspended), allowing the tutor to easily keep this information. The UI is interactive and allows the drag-and-drop of all the articles. The software encourages tutors to follow ABA rules, proposing the correct trial sequence, but it is also flexible, allowing the tutor to modify the action to be performed. Pushing the level button activates the trial on the child's device. At the end of the session, the tutor can leave additional comments about the session to communicate information to the other tutors, and then close the session.

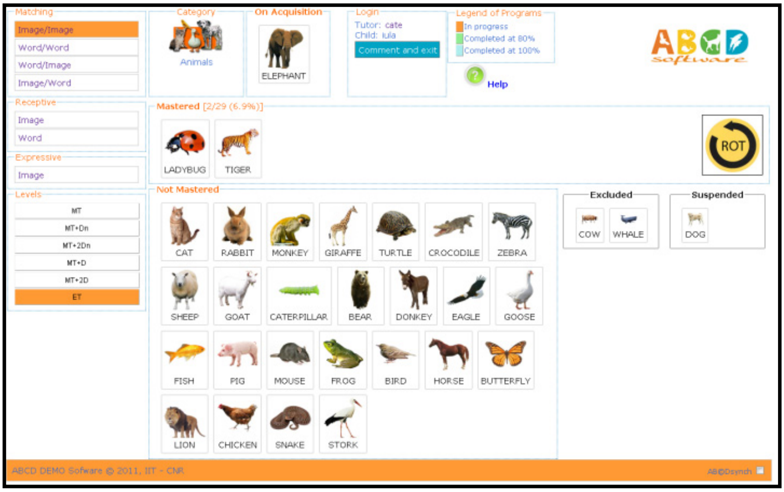


Fig. 2. The Tutor User Interface for setting the trial

When the tutor starts a session with a child, the history of previous sessions is extremely valuable for setting up the current work. In traditionally performed ABA sessions, access to this information involves reading an enormous amount of records on paper, requiring time and effort. To aid tutors in this operation, we created an interface, visible immediately after the login phase, that provides data about previous sessions with different levels of detail. The page shows an ordered list of links labeled with the date and name of the tutor who performed the session. Clicking on a link shows data related to each program performed in that session; a second click on each level summarizes the statistics related to the single trial.

3.3 The Child's User Interface

The child's user interface is available on the touch screen device depending on what was selected by the tutor. ASD children's learning and communication development are often supported by the visual channel, which is usually the most effective [5]. For this reason, the kind of pictures used was carefully considered during the design phase, and the pictures chosen are photos in a realistic environment, without background. The window on the screen is organized to make the task clear to the child, avoiding confusion. For example, in matching programs the target element is placed at the bottom, closer to the child, and it is the only element of the UI that can be moved. To avoid any errors, respecting the ABA principle of errorless learning, if the child tries to match the target element with an incorrect item, the target is rejected and returns to the initial position; otherwise, if the match is correct the target element is attracted by the correct article, thanks to a region of attraction placed around it. All actions -- dragging, touching -- performed by the child are recorded in the database in real-time and instantly available on the tutor's panel, so the tutor can monitor what is happening on the child's screen and obtain a summary of the trial performance of that level. In case of Matching programs, since the software does not allow incorrect matching, the system records each child's attempt to match an incorrect item as a child's error.

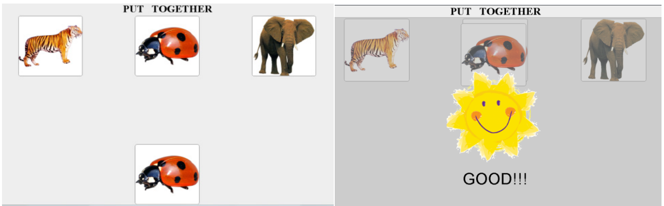


Fig. 3. The child's UI: matching trial with two distracters -- feedback after a correct trial

Once the trial is activated, the Child environment is set up according to this trial and the child's profile. In Fig. 3, the system proposes a trial for which the child has to match the ladybug. Since the child is non-receptive, i.e., does not understand vocal instruction, the canvas supports the child by proposing the command as a label on top of the screen. It is important to note that this kind of aid is consistent even when the child is not yet able to read because when working with non-receptive subjects, the tutor utilized a label with the written command (e.g.: 'put together' as in Fig. 3) and due to the intensive repetition of the exercises, the child learns the label associated to commands. According to the importance of reinforcing a correct behavior, each correct trial is rewarded with a little animated gif that appears on the screen (Fig. 3, right).

When the child's screen is triggered, the tutor UI is overlapped with a second UI for evaluating the trial. This UI is a sort of a tutor control panel and reports information on the trial displayed on the child's device -- first row in the upper part of Fig. 3 -- and the real-time summary of statistics regarding the child's performance on the trials performed. This is extremely helpful for the tutor, who can obtain an overall view of the exercises. There are the sums of trials for level of prompt provided: 0, 20, 50, 80, 100%. The child's behaviors are recorded (error, non-collaboration, self-stimulation), as well as any tutor errors. If the child does the exercise independently without prompts, or if the tutor considers that the child has done the best he/she could do, an interactive reinforcement such as a video, an animation, etc., can be activated on the child's user interface. It is possible to activate the reset of the child's user interface, if the proposed selection is not suitable for the child. In the end, the last three rows provide the tutor with a summary of the child's performance on the trial of this level/program, with the total number and percentage of correct, incorrect and prompted trials. If a touchscreen tablet is not available, the tutor and the child can share the same computer. The tutor has to deselect the check box in the lower right part of the control interface (i.e., AB©Dsynch in Fig. 2).

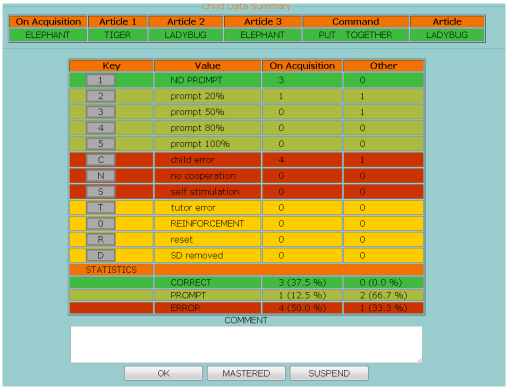


Fig. 4. The tutor control panel

3.4 Synchronization between the Child and Tutor UIs

In order to verify the best kind of touchscreen devices to use with ABCD SW, we tried two different large touchscreen monitors and five types of touchscreen tablets, actively involving the children in testing the devices. The use of big touchscreen monitors was quickly discarded because they are not very sensitive, a quality necessary for providing ready feedback to the child's actions. In addition, this solution needed USB cables to connect a PC and the additional monitor. Since autistic children are often attracted to cables, their presence could be a distraction; or worse, they could be used improperly, causing safety issues. Another advantage of the touchscreen tablet is the system's portability, which makes it possible to perform an ABA session anywhere, anytime.

4 The Pilot Test

A user test with seven children was carried out during the 2011-2012 school year to verify the children's acceptance and pleasure, the efficiency of the tool for tutors and the satisfaction of families. The main goal of this test was to refine and set up the ABCD SW, enhance its functions and improve the user interfaces.

ABA intervention requires a high level of attention from the children, so the first 4 months were spent bringing all children up to the same level in the ABA methodology, since three out of seven of the enrolled children had already received intensive ABA intervention for more than 2 years. Children were recruited by an association promoting ABA for the care of children with autism, that contacted parents and set up specific agreements with the children's families. Intentionally, the age of the selected sample was spread out from 2.5 to 6 years (6 subjects), with one 10-year-old child included in order to expand our observations about the range of possible engagement of the software by different ages. All the children are low-functioning. An intensive ABA intervention was carried out with an ABA team 15 h/week in the afternoons and from 10-14 h/week at school, ensuring at least 25 h/week.

We observed the seven children using the software for 4.5 months, also collecting the ABA tutors' observations and requests, crucial for improving the software and better adapting it to their and the children's needs. The data collected in the ABA sessions by means of the ABCD SW is enormous and is still under evaluation by ABA senior tutors and a psychologist. However, preliminary data on children's reactions collected by the tutors and the children's caregivers are encouraging; all the children were attracted by the didactic software and were happy to work on the device, with an accompanying increase in positive behavior.

4.1 Software Improvements and Lessons Learned

Several adjustments and new functions were added during the pilot test. First of all, we realized that mastery of an article (the goal of the ABA programs described here) is not a definitive process. Sometimes it happens that a child appears to be unfamiliar with an article previously considered mastered, so that the work with this article must be repeated. For that reason we completely changed the original schema by introducing a different status for each article, which can change over time: "acquisition", "mastered", "suspended". According to that new schema, a mastered article can return to acquisition status when required. For similar reasons, if an article appears to cause behavioral problems or seems to be very difficult for the child, the work should be diversified and the obstacle bypassed. The proposed solution was to allow the tutor to categorize an article as "suspended" permitting a later re-approach (Fig. 1). In addition, the tutor can decide a priori to exclude some articles if they consider them not suitable for a specific child.

Another problem concerned the possibility of adding comments as a final step during a session. Entering data manually requires time and attention, not always available when working with ASD children, e.g., the child may initiate self-stimulation or other problem behavior requiring the tutor's attention. In order to address this problem, we implement the possibility of updating this type of data offline, protecting the integrity and consistency of global session data.

The software favors easy configuration of the most suitable child UI because it proposes the programs/levels and single trials according to a set of rules. As an example: a different canvas for the child's screen is proposed according to the child's characteristics, since the command can be presented or not depending on whether the child is non-receptive or receptive. However, especially in the case of the program "Receptive word", an image of the target article appears near the stimulus (Fig. 5); for some children that kind of discriminative stimulus can serve as a little prompt so the tutor can control its appearance or disappearance with a simple key.

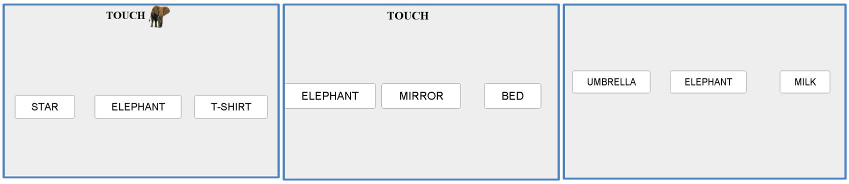


Fig. 5. Example of software customization: different UI for Receptive program

Concerning personalization of software, a new level of customization introduced was the possibility of loading personal items in some categories. For example, this is the case of the *Places* (of the house and school) and *Family* categories that obviously contain different images for each child. In this case the tutor/parent can add content (images) to the database, associating them with his/her child thanks to a Drupal interface designed for this purpose. The association with the child user profile will make this content invisible for the other users.

Summarizing, the main characteristics of the ABCD SW customizations depending on a child's abilities are as follows.

Customizable features:

- Personalization of multimedia reinforcements (child configuration file)
- Personalization of the categories Relatives and Places (UI for items upload)
- Exclusion or suspension of problematic items (drag & drop in the tutor UI)
- Fading of the visual support for the trial ('D' Key)
- Reset of a trial if inadequate for the child ('R' key)
- Child age < 5 years: Matching programs with words are removed from the tutor UI (child configuration file)
- If the child is Non-Receptive, the command is shown as a word in the trials (child configuration file)
- If the child is Non-Expressive, and thus cannot answer vocally, the UI shows 3 words from among which the child will choose the right one (child configuration file).

5 Usability: A Qualitative Assessment

An initial subjective evaluation of the software's usability was carried out through an online survey proposed to the ABA teams following the participants in the user test phase. The survey was created using Google Docs, sending an email request to potential participants. The questionnaire comprised two parts aimed at evaluating both the classic ABA intervention and that supported by computer. In the following, the main results concerning the usability of the tool are reported. A five-element Likert scale has been used as answer range and one option has been added if the participant feels unable to provide an answer. Forty-seven answers were collected, according to the distribution showed in Fig. 6. Thirty-three of the 47 (70% of the sample) are ABA professionals with specific training (tutors, senior tutors and ABA consultants). According to the ISO definition of usability, we asked participants for efficacy, rapidity and child satisfaction in the use of the software. Answers clearly acknowledge the validity of using ABCD SW as a support for complementing and enhancing the ABA intervention.

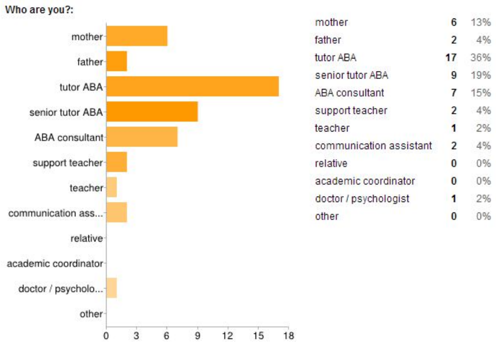


Fig. 6. Distribution of Survey Participants

In general, children appreciated the use of the touchscreen tablet (Fig. 7). One of the seven subjects in particular, who was unable to use a computer or watch TV, showed great interest in and increased attention to using his personal learning environment on the touchscreen tablet.

Fig. 7. Children's feelings about the software

Concerning the usefulness of the software, 58% of the sample acknowledged that it is useful (28%) or very useful (30%) while only two persons did not agree (Fig. 8).

Fig. 8. Usefulness of the software

Concerning the ease of use of the software, only one person found it difficult while 26% declared that it was simple and 34% very simple, making a total of 60% who perceived the system to be easy to use as shown in Fig. 9.

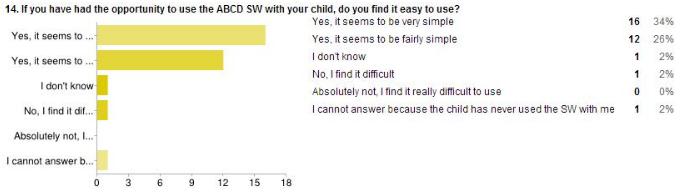


Fig. 9. Ease of use of the software

Another fundamental aspect concerns the rapidity of the ABA software compared to classic ABA intervention. Of the total sample 55% rated the software as efficient (38%) or very efficient (17%), while two persons were not satisfied (i.e., believed the software is not very efficient) as shown in Fig. 10. However it is important to note the lack of an ADSL connection at one child's home; he was forced to use a mobile connection, with notable degradation of performance, which impacted on the satisfaction of the child and the ABA team. A guaranteed bandwidth is needed in order to take advantage of the Web architecture. Obviously, this delay can also impact on the evaluation of the usefulness of the software.

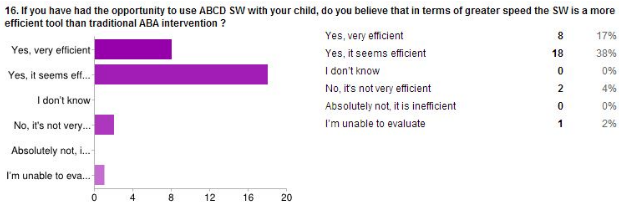


Fig. 10. Speed of the software vs traditional ABA intervention

6 Conclusions and Future Work

Accessible education for children with learning disabilities is still a great challenge and one of the main priorities of our society. In this paper we discussed the architecture of a Web application called ABCD SW, based on ABA, AAC and DTT, and created for teaching low-functioning children with autism. The software supports tasks performed by tutors during an ABA intervention, so they can automate the trial set-up and the definition of the sequence of proposed trials, enabling evaluation of trials executed by the child. In this way, 1) the sequence of trials can be proposed more or less rapidly, according the subject's learning pace; 2) data gathering is accurate since tutors are forced to provide an evaluation in order to continue the intervention, and 3) the population of categories is scalable -- a web interface allows the upload of new items in the category -- since no change to the code is required to add new elements.

Thanks to the wide diffusion of Wi-Fi networks, ABCD SW can be also used with two mobile devices (a laptop and a touchscreen tablet) in places different from school and home, such as parks, facilitating the generalization of a mastered article in different places. Furthermore, also due to the CMS support, the software is multilingual: all the texts of the user interfaces, the labels for articles or categories, and the commands of the exercise can easily be shown in different languages (with no changes to software code, only the translation of these items is required).

An initial subjective evaluation of the software confirmed its value and potential as a complementary approach for teaching autistic children with severe attention impairment.

Current efforts focus on creating a tool for efficient and accessible visualization of the child's learning progress and trends over time by means of graphics. In future studies it would be extremely interesting to investigate the true potential of the children. At the moment, ABA intervention requires one-to-one interaction with specialized personnel and is very expensive; thus it is limited to a few daily sessions. This is a limitation, and we do not completely exploit the actual learning abilities of the children. The creation of a controlled environment that allows children to progress in the hierarchical levels alone without the supervision of a tutor once a degree of success is reached might be a good solution. The children's controlled zone implemented in the ABCD SW is the first step toward maintaining the errorless principle. Furthermore, in low-functioning children with autism, mastered items are unstable -- if not continuously refreshed over time they may be lost. The idea is to develop a safe environment for the child to perform the going-over phase on mastered items, alone. It is important that the child use the tablet only as a learning tool and not as an auto-stimulatory activity. New studies are needed to investigate this potential as well as explore new strategies for teaching autistic children. The Internet and Web-based architecture help us provide simple tools that are available anywhere, anytime as an aid to learning.

Acknowledgements. We thank the Regione Toscana, which funded this project within the framework of the "FAS 2007 2013 Delibera CIPE 166/2007 PAR FAS Regione Toscana Action Line 1.1.a.3" (14 Feb 2011-14 Ago 2013). We also thank the children who participated in the pilot test, their families and the entire ABA team.

References

[1] Buzzi, M.C., Buzzi, M., Gazzé, D., Senette, C., Tesconi, M.: ABCD SW: Autistic Behavior & Computer-based Didactic software. In: Proc. International Cross-Disciplinary Conference on Web Accessibility (W4A 2012), vol. 28, 2 pages. ACM, New York (2012)

[2] De Leo, G., Leroy, G.: Smartphones to facilitate communication and improve social skills of children with severe autism spectrum disorder: Special education teachers as proxies. In: Proceedings of IDC 2008, pp. 45–48 (2008)

[3] Hayes, G.R., Hirano, S., Marcu, G., Monibi, M., Nguyen, D.H., Yeganyan, M.: Interactive visual supports for children with autism. Personal and Ubiquitous Computing 14(7), 663–680 (2010)

[4] Hirano, S., Yeganyan, M.T., Marcu, G., Nguyen, D.H., Boyd, L.A., Hayes, G.R.: vSked: Evaluation of a System to Support Classroom Activities for Children with Autism. In: Proceedings of CHI 2010, pp. 1633–1642 (2010)
[5] Hodgdon, L.A.: Visual strategies for improving communication: Practical supports for school and home. QuirkRoberts Pub., Troy (1995)
[6] Kivlehan, D., Melvin, H.: Development of a Web-Based Technology Engine for ABA Delivery (2008)
[7] Monibi, M., Hayes, G.R.: Mocotos: Mobile Communications Tools for Children with Special Needs. In: Proceedings of IDC 2008, pp. 121–124 (2008)
[8] Rosenwasser, B., Axelrod, S.: The contribution of applied behavior analysis to the education of people with autism. Behavior Modification 25(5), 671–677 (2001)
[9] Rosenwasser, B., Axelrod, S.: More contributions of applied behavior analysis to the education of people with autism. Behavior Modification 26(1), 3–8 (2002)
[10] Russo, D.C., Koegel, R.L., Lovass, O.I.: A Comparison of Human Automated Instruction for Children with Autism. Journal of Abnormal Child Psychology 6(2), 189–201 (1978)
[11] Swettenham, J.: Can children with autism be taught to understand false belief using computers? Journal of Child Psychology and Psychiatry 37(2), 157–165 (1996)
[12] Weiss, M.J.: Expanding ABA intervention in intensive programs for children with autism: The inclusion of natural environment training and fluency based instruction. The Behavioural Analyst 2(3), 182–186 (2001)
[13] Whalen, C., Liden, L., Ingersoll, B., Dallaire, E., Liden, S.: Behavioral Improvements Associated with Computer-Assisted Instruction for Children with Developmental Disabilities. The Journal of Speech and Language Pathology - Applied Behavior Analysis 1(1), 11–26 (2006)

Pervasive Interventions to Increase Pro-environmental Awareness, Consciousness, and Learning at the Workplace

Dirk Börner, Marco Kalz, Stefaan Ternier, and Marcus Specht

Centre for Learning Sciences and Technologies, Open Universiteit, Heerlen, The Netherlands
{dirk.boerner,marco.kalz,stefaan.ternier,marcus.specht}@ou.nl

Abstract. This paper reports about pervasive interventions at a university campus to increase the pro-environmental awareness, consciousness, and learning of employees. Based on an assessment of the research gaps in this problem area we present results and design implications from three intervention iterations. While in the first intervention the focus was on increasing awareness through information distribution with ambient learning displays on the campus, the second iteration provided personalised feedback to employees with the help of a sensor network and different client applications. The third iteration then implemented a game-based learning concept. Results reveal that these approaches are effective on different levels and that a combination of these elements can lead to increased pro-environmental consciousness, learning and hopefully a sustained behaviour change of employees.

Keywords: Environmental Learning, Pervasive Education, Ambient Learning Displays, Game-based Learning, Design-based Research.

1 Introduction

While energy conservation in the domestic context has been sufficiently studied and evaluated, there is only a small amount of studies focusing on energy conservation at the workplace. A study by Siero et al. [1] showed that the offering of information and learning opportunities about pro-environmental behaviour has the potential to change the attitude and behaviour of employees. Nonetheless, the workplace context is different from the domestic one. A recent study by Lo, Peters, and Kok [2] revealed that the main differences are that the costs of energy consumption are not monitored nor paid by the employee and that the organisation's structure, size, goals etc. has an influence on individual behaviour. Furthermore the authors stressed the importance to understand the psychosocial determinants of pro-environmental behaviour at the workplace, which differs from the domestic context. In conclusion the authors presented a framework consisting of individual and organisational determinants that can influence individual behaviour at the workplace and finally identified five factors: attitude, awareness, self-efficacy, subjective norms, and habits. Kollmuss and Agyeman [3] presented a complex model of pro-environmental behaviour that integrates internal factors such as personality traits or environmental consciousness and external factors

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 57–70, 2013.

such as infrastructure or political context. Additionally they investigated and incorporated possible barriers to pro-environmental behaviour. These barriers are mainly responsible for the gap between attitude and action, also referred to as engagement gap. Among others the identified barriers were lack of environmental consciousness and knowledge, negative or insufficient feedback about behaviour, as well as missing internal and external incentives.

We have recently conducted a study that has shown that only 25% of employees in an academic organisation are concerned about the financial consequences of their individual consumption for the organization [4], while financial incentives are currently one of the major driving forces in the domestic context. Besides that the surveyed employees reported that they feel unaware about the organisational and individual energy consumption and respective conservation possibilities. Thus around one third requested more detailed information as well as clearer incentives from the employing organisation.

Foster et al. [5] approached the problem from an interaction-design perspective. In focus group sessions with stakeholders from different levels of an organisation they have identified the problem as consisting of motivational, social, organisational, and technical issues. Furthermore they identified a "research knowledge gap present in understanding the end-users of energy in the workplace and, therefore, the design of appropriate and achievable workplace energy interventions, particularly those that encompass novel ways of encouraging people to adopt positive energy usage behaviour whilst at work." This formulated research gap is exactly the purpose of the research reported in this article. In this paper we report about three connected interventions that had the target to increase awareness, pro-environmental consciousness, knowledge and last but not least behaviour change of employees of an academic institution. By this approach we target the internal factors and external barriers of the model by Kollmuss and Agyeman [3], namely on the lack of knowledge and environmental consciousness and the lack of incentives and insufficient feedback about behaviour.

The paper is organized as follows. In the next section we briefly present design-based research as our methodological basis for the interventions and present the problem context. In section 3 we present the design propositions, narrative and results of an intervention that had the target to increase pro-environmental awareness and knowledge about energy conservation amongst employees with ambient learning displays. In section 4 we report about an intervention in which energy consumption sensors have been deployed in combination with personalised consumption information. In section 5 we report about a recent intervention that has been informed by results of the first and second intervention. In this section we introduce the design propositions, narrative, and results of a game-based intervention that combined several technologies and incentive mechanisms to foster pro-environmental consciousness and to initiate behaviour change among employees. In section 6 we discuss the results of the iterations and summarise their value for future conservation campaigns and activities in organisations and environmental learning. Last but not least we discuss the limitations of the approach taken and provide a perspective for future research.

2 Method and Context

Our research is based on the design-based research methodology [6]. Design-based research addresses complex problems in authentic environments, integrates design principles with technological affordances, and conducts reflective inquiry with the target to refine learning environments and to identify new and emerging design principles. In contrast to predictive research that is targeting on the specification of new hypotheses design-based research is targeting in the constant refinement of problems, solutions, methods and design principles [7]. According to the design-based research collective [8] the goals of developing theories and designing learning environments are intertwined and these activities constantly inform each other. Another goal of design-based research interventions is to communicate relevant implications to practitioners and other educational designers. In this sense, we see our contribution as an input to other practitioners, institutions and educational designers that want to increase the environmental consciousness and foster conservation at the workplace.

In our case the context of the research stems from a long-term national agreement on energy efficiency that public institutions have with governmental agencies. In this agreement the university agreed on reducing the energy consumption by 2% each year until 2020 and to raise awareness on this topic among employees. Before we started our intervention awareness raising was limited mostly to some stickers and posters and employees were not aware about this institutional goal. Therefore we have conducted research in three iterations among the employees working at the main university campus. There has been only a small overlap between the participants of the single iterations. So all three iterations explored the potential of different interventions on the lack of knowledge, environmental consciousness and the lack of incentives and insufficient feedback about behaviour. The feedback collected from one iteration has informed the design of the following iteration.

In a first design iteration we have solely focused on increasing awareness and knowledge about pro-environmental behaviour by utilising public displays on the campus. The main question here was if ambient learning displays are suitable means to increase environmental awareness and knowledge of employees. To be able to inform future design activities that employ public displays for this purpose we have designed four display prototypes that have been evaluated in different buildings in the campus. The second iteration has taken up feedback by employees and findings from related studies showing that personalised feedback is an effective way of encouraging pro-environmental behaviour. Hence we have deployed personalised consumption statistics via a sensor network as well as desktop and mobile client applications.

Finally in a third iteration we aimed to go beyond a pure information presentation and combine different components in a game-based learning intervention to increase the involvement of employees with challenges, rewards, and more complex problem solving. In the following sections we present the respective design propositions, design narratives, data and analyses, and finally discuss how our findings influence future interventions and design decisions.

3 First Iteration: Employing Public Displays to Increase Environmental Awareness

3.1 Research Questions and Design

In the first iteration of our intervention we have focused on the use of public displays on the campus to increase the awareness for pro-environmental behaviour and energy saving potential. This project is related to an on-going PhD project about the design and application of ambient learning displays for the situated support of informal and non-formal learning scenarios in ubiquitous learning environments by enabling learners to view, access, and interact with contextualised digital content presented in an ambient way. Based on an earlier conceptual framework [9] and a literature review [10] that has analysed and classified the existing prototypes we have focused on the question how ambient learning displays need to be designed to serve as suitable means to go beyond the static presentation of energy conservation information as it was prior to our initiative. For this purpose we used an existing model depicting the design dimensions of ambient systems by Pousman and Stasko [11] with its components information capacity, notification level, representational fidelity, and aesthetic emphasis. To date, the influence of these dimensions on environmental learning and awareness had not been examined. Based on our analysis of the state of the art we decided to not focus on the components information capacity and aesthetic emphasis. Information capacity is determined by the amount of information represented by the system, which needed to be consistent to measure a learning outcome reliably. The emphasis put on aesthetic is a highly subjective measure that heavily depends on the context in which the ambient system is used. Therefore this dimension was out-of-scope for examining a prototypical system. Consequently we employed the design dimensions representational fidelity and notification level to influence our intervention. The representational fidelity dimension describes how the data is encoded on an ambient system while the notification level depicts the degree of user interruption evoked by the displays. Both dimensions covered a broad design spectrum ranging from indexical to symbolic representations and blind to interruptive levels of notification. Consequently the respective extremes were used to manifest our four design prototypes for the intervention. The possible combinations of the representational fidelity and notification level design dimension resulted in four groups covering all different combinations.

The resulted interventions have been implemented in four buildings on the campus. These four design prototypes were manifestations of the possible combinations of components, namely (1) blind notification and indexical representation, (2) blind notification and symbolic representation, (3) interruptive notification and indexical representation, or (4) interruptive notification and symbolic representation. Since the focus of the intervention was to analyse effects of ambient displays that are not in constant focus of attention employees were not asked directly to participate in the experiment and watch out for the treatment. After the measurement of a baseline display prototypes were installed silently in the entrance area of the four office buildings. The prototypes consisted of a Dell M2010 notebook with built-in speakers and webcam but without attached keyboard or mouse. The speakers were used to send out audio notifications, while the webcam was used to enhance the functionality of the notebook with a custom-built movement/attention sensor. The sensor was built using the Processing[1] development environment and the open source computer vision library for Processing. Slideshows were used on the prototypes that consisted of three parts: information regarding energy consumption in the building, generic saving tips, and the overall conservation potential. The most important information on each slide was highlighted in red and contextual information was highlighted in blue. The first part contained information showing the average electricity consumption per working day of each employee, the whole campus, and the building the display was located in.

The prototype variation on notification level was implemented using the custom-built movement/attention sensor to trigger the notification as well as the built-in speakers to play back a respective audio file. For the interruptive treatments one audio notification was played when the sensor detected movement and another one when the sensor detected that someone turned towards the display. For blind treatments any notification was omitted. The variation on representational fidelity was implemented as two distinct means of information presentation. For the indexical representation raw data facts were used to communicate consumption information, saving tips, and conservation potentials. In contrast, topic-related icons were used for the symbolic representation of the data, e.g. light bulb icons representing 5W each.

3.2 Data and Analysis

A total of 563 university employees were asked to participate in the study. 190 employees responded to the pre-test. 101 employees responded also to the post-test. The university's campus consists of four main buildings, namely "Athabasca", "Chiba", "Madrid", and "Milton Keynes". Only employees working in one of these main buildings were considered as participants of the study. For the purpose of the study the participants were then divided into groups depending on the building they are working in. The resulting four groups (N=94) have been subject of our intervention. The participants (37 females and 57 males) were in the age range between 26 and 65 and have been working for the university for at least two and up to 26 years. These participants confirmed in the post-test that they have recognized the prototypes. Out of the 94 participants 12 participants were exposed to the prototype with blind notification and symbolic representation, 35 to the prototype with the blind notification and indexical representation treatment, 12 to the prototype with the interruptive notification and indexical representation treatment, and 35 to the prototype with the interruptive notification and symbolic representation treatment.

An online-survey application was used to deliver questions to the participants. The survey-instrument consisted of four blocks of questions. The first block was related to the actual knowledge about consumption, one block of questions was related to the confidence to estimate individual and institutional consumption and conservation potentials, one block of questions was related to the awareness need and estimated

---

[1] `http://processing.org`

effectiveness of higher awareness, and the fourth block of questions was related to environmental concern and conservational attitude. The questionnaire was not anonymised to allow pairwise-comparison of pre-test and post-test data. Some open questions were asked to collect also general comments and feedback to the intervention. Most questions consisted of 5-point or 7-point Likert-type scale items giving choices ranging from 1 (not at all) to 5 or 7 (completely). Others requested multiple-choice or open answers. The post-test questionnaire then also contained questions related to the individual perception of the ambient learning display and comprehension aspects. The construct environmental learning outcome was calculated by summing the individual component gains, i.e. knowledge, confidence, awareness, and concern. The pro-environmental behaviour was determined by the conservation activities performed as well as the actual energy consumption data. Thereby performed conservation activities were also measured directly within the introduced questionnaire. The respective part simply asked for the number of activities performed. Finally the actual energy consumption data was obtained from the institutional facility management system on a daily basis.

3.3 Results and Discussion

Analysing the results of the pre-test and post-test data shows the following results:

- The group with interruptive notification and symbolic representation had the largest gain within the construct environmental learning ($M = 0.90$, $SD = 1.56$) and the group with interruptive notification and indexical representation the smallest ($M = -0.17$, $SD = 2.14$);
- The largest knowledge gain could be measured for the group with blind notification and indexical representation ($M = 0.31$, $SD = 0.76$);
- The largest confidence gain ($M = 0.21$, $SD = 0.83$) as well as the largest awareness gain ($M = 0.49$, $SD = 0.91$) could be measured for the group with interruptive notification and symbolic representation;
- The largest concern gain could be measured for the group with blind notification and symbolic representation ($M = 0.22$, $SD = 0.98$).

A Kruskall-Wallis test was conducted to explore the influence of the different treatment conditions on the environmental learning outcome as well as the individual component gains. But none of the effects were significant to demonstrate the superiority of the one prototype design against the other. Across all groups the comparison of the single means with a number of Wilcoxon signed-rank tests shows that the deployed prototypes significantly influence awareness and knowledge. In total participants scored significantly better on the knowledge component after the treatment ($M_{\text{Total}} = 5.07$, $SD = 0.72$) than before the treatment ($M_{\text{Total}} = 5.29$, $SD = 0.54$), $z = -2.60$, $p = .009$, $r = -.19$. The effect size is small. Participants felt a significant lower awareness need after the treatment ($M_{\text{Total}} = 4.29$, $SD = 1.34$) than before ($M_{\text{Total}} = 4.65$, $SD = 1.39$), $z = -3.33$, $p = .001$, $r = -.24$. The effect size is again small. This reveals that the deployed prototypes help to examine and comprehend and lower the awareness need of employees. But the qualitative results of the post-test also show that there is a need for alternative ways to motivate employees to save energy at the workplace, as for instances clear incentives are missed or the provided information was to generic. In addition, research shows [12] that interventions that are only based on the implementation of ambient displays might lead to an increased comprehension of the display and represented information, but at the same time the user interest can decrease. This could also lead to less commitment of employees. These results have led us to a second iteration that evaluates the potential for personalised feedback on energy consumption of employees. The design and results of this intervention are presented in the next section.

4 Second Iteration: Personalised Feedback to Increase Environmental Consciousness and Knowledge

4.1 Research Questions and Design

While the use of ambient learning displays might lead to increased awareness, one of the problems of these technologies is the low level of personalisation. In fact, employees that have participated in the first intervention have reported the need to receive personalised information about their energy consumption at the workplace. We have taken the feedback variables by Mory [13] as guiding framework for the design of the personalised energy consumption feedback, The main research question for this iteration was how we can design effective personalised feedback to inform employees about energy consumption of their working environment. This iteration is based on related work about the design of so called eco-visualisations [14].

To identify the most effective way of providing personalised feedback to employees we deployed energy consumption sensors in parts of a campus building. This deployment included 7 classic office rooms, 7 hot desk rooms and two meetings rooms. For the hot-desk setup a flexible way of registering employees at the workplace has been developed. The sensor network allowed the reliable sensoring and logging of consumption information on several aggregation levels. The sensor network was deployed in a way that energy consumption statistics could be aggregated on an individual level, a room level, or an appliance level. With this logging infrastructure in the background we have developed mobile, desktop, and web client applications to allow employees a ubiquitous access to these information and to be able to reuse the information in the organisation as flexibly as possible. In addition this setup eventually enables the comparison between e.g. colleagues, research groups, or departments.

4.2 Data and Analysis

The data collection and analysis has been conducted on three different levels. As a basic level the current environmental awareness of employees was assessed. The subgroup of employees who used the web/desktop client received again the same questionnaire after the intervention. In addition the participants evaluated different display designs for a mobile application. The baseline measurement was delivered via an online-survey to employees in the building where the intervention should be deployed later. Respondents (*N*=58) were asked questions about their awareness, concern, and attitude regarding energy consumption and conservation at the workplace. After deploying the prototype the study has been repeated among the employees who actually used the prototype (*N*=14). Results were treated as categorical data and statistically analysed using medians. In total three different mobile display designs have been evaluated by the participants (*N*=17). The first design used indexical representations of real-time usage and consumption data with dashboards and graph visualisations. The second design presented a polar bear icon in an environment that adapts to the current power usage, ranging in six steps, i.e. from lots of ice, food and bears for low usage to one bear or even an empty sea for high usage. The third design used a symbolic colour spectrum from blue for low to red for high consumption and additionally textual level from 1 to 10 for presentation. In individual sessions the designs were presented and explained to participants, who were already familiar with the infrastructure and the developed applications.

4.3 Results and Discussion

A full report about this iteration is given in [4]. Interestingly the deployment of the feedback intervention in the subgroup has made them aware that they are not active enough and need more information and knowledge about energy conservation at the workplace. In addition most participants of this intervention communicated the need for incentives to save energy. Results show that although the display prototypes have not been used extensively the information presented was perceived well and understood. Information granularity of the visualization has satisfied the needs of employees.

In the design evaluation of the three display prototypes the third display design received the most positive responses. The qualitative comments showed that the second design (polar bear) was perceived as too “childish” for the workplace context. Participants recommended the metaphor to be used as a group or community representation but not for personalised feedback. Also the logic of the metaphor has been criticized for example with regards to using solar power. One participant also came up with the idea, to use personal images for the feedback instead. These images could then be toned with hot (orange/red) or cold colours (blue/green) for high/low usage and be shown on a personal device. These results have led us to a third iteration. This time we have not only focused on awareness and feedback, but also on environmental consciousness, incentive mechanisms, and gamification approaches. The design and results of this intervention are presented in the next section.

5 Third Iteration: Game-Based Learning to Increase Environmental Consciousness and Behaviour Change

5.1 Questions and Design

For the third iteration we had the target to go beyond increasing awareness and knowledge. Instead we wanted to focus in addition on the lack of incentives and the lack of feedback identified in the model by Kollmuss and Agyeman [3]. It has been shown that projects that focus only on behavioural approaches like operant conditioning, incentives, or rewards are effective in short-term but not effective in long-term energy consumption behaviour [15]. According to the authors, the effects diminish from the moment the reward is not given anymore. Therefore our target in the project was to combine behavioural approaches with a motivational and social influence approach. Thus, our idea was on the one hand to combine motivation approaches with different barriers and factors in the model (see Table 1).

Table 1. Identified problem areas and game elements

Environmental Consciousness	Lack of incentives	Lack of feedback
• Knowledge components	• Digital Badges	• Testing
• Energy statistics	• Rewards	• Challenges

To connect the different components we used gamification approaches and game elements as the framework for the components. Gamification is the concept of applying game-design thinking to non-game applications to make them more fun and engaging. Werbach & Hunter [16] differentiated between gamification approaches that focus on an organizational benefit and approaches that focus on a personal benefit. Our initiative is focused on an individual behaviour change with clear benefits for the organisation. On the other hand there is a secondary individual benefit of our initiative since future generations might benefit from pro-environmental behaviour that happens right now.

The main design challenge for the third iteration was to define a challenging but also realistic game-based learning scenario that employees are able to follow during their working time and to provide appropriate and efficient incentive mechanisms that keep the participants motivated. For these two layers of the intervention we have used and integrated existing technology, namely the ARLearn application platform [17] to design and implement the scenario as well as the Mozilla Open Badge infrastructure [18] to provide incentives. As alternative reward to the digital badges we have provided weekly small prizes to participants of the game, such as book voucher for the employee who collected the most information, activity voucher for the most active employee who performed all the actions, electronic media voucher for the employee who mastered all the challenges. If there was more than one employee qualified for the prize then the winner was chosen at random. Furthermore there was an overall prize for the best player (aka. the greenest employee), announced and awarded after the game.

The game design was constructed from the following game elements: information, video, action, challenge, activity, quiz, and badge. The information element provided the users with all the important knowledge, e.g. about the game, energy consumption details, conservation possibilities, saving potentials etc. As a variation of the information element videos provided the users with simple tips on how to conserve energy. Thereby we made use of available topic-related material.

Action elements were used to get users active and let them do something, among others to find something out, save some energy, explore the campus, etc. To perform actions they had to leave their workplace and reach different places on the campus, e.g. the game flags we deployed in the centre of the campus. Most of the time actions combined information clues and assignments at the same time. A sample action looked like this: “Athabasca is a rather small building on our campus, which consumed in total 1154 kWh electricity last week and 200 kWh on average per working day. With 256 kWh the highest electricity consumption in Athabasca was on Thursday. Last weekend Athabasca consumed 152 kWh without anyone in the office. Now look for the small QR code attached to the 'Chiba' flag pole and scan it.” Challenges invited the users to elaborate and reflect, e.g. by sharing their opinion and personal experience etc. These items were about the users, their ideas, opinions, and experience about them and their workplace using different kind of media. A sample challenge looked like this: “The Mindergie game is about YOU so we would be happy to know: What are your reasons to participate in the game? To do that simply press Provide Answer, record an audio statement, and publish it. You can record more than one statement if you like. When you are finished, please go back to the list to continue.”

The activity element was introduced to allow users to register their conservation activities. The idea was to get an impression on their habits, so they were asked to be honest and only register activities they had really done. Following that codex they were allowed to register as many activities as they liked from a list that was adapted weekly to the theme of the week, e.g. switch off appliances instead of leaving them on stand-by.

The quiz element was mainly used to assess the knowledge acquired during the game, e.g. by reading all available information or watching the information videos. Usually this element became available only after accessing all necessary elements. The outcome was taken as basis to issue badges.

Finally when users demonstrated a skill, achievement, or quality during the game they were usually rewarded with a badge. The respective element then became available and could be used to store the earned badge in the personal backpack. A set of badges has been designed for the project. In total four types of badges were used, one for the general gameplay and one for each category. The different types of badges are distinguished by form and colour. Each badge is characterized by a unique symbol illustrating it’s meaning. Furthermore each badge can have three different states or levels reaching from bronze over silver to gold.

5.2 Data and Analysis

All employees were asked to register for the game and become the greenest employee of the university. As the game is based on ARLearn, the only requirement was to have an Android smartphone or tablet available as well as own a Google account. Furthermore as the game uses Mozilla Open Badges it was necessary to register with Mozilla to be able to keep track of all accomplishments also beyond the game. We had a limited amount of Android devices available to borrow. After registration the participants were invited to participate for the next 4 weeks in the weekly game rounds. The designed game was about energy consumption and conservation at the workplace. The main purpose was to provide the users with useful information on the What, Why, and

How as well as to motivate them to get involved and committed. The game was played in weekly rounds structured around certain topics. The first week started really simple with an introduction to the game and the technologies used to play it. The second and third week were then mainly about electricity and respectively gas consumption and conservation at the workplace. Finally the fourth week dealt with individual consumption footprints and alternative conservation strategies.

The main items that could be found in the rounds were: information, actions, and challenges. All the items appeared in the ARLearn message list or opened automatically when they became available. Each item type was noted (in brackets) in front of its title. From time to time users were asked to answer questions, either as part of an item or in the course of quizzes. Usually when answering questions, read information, perform actions, or master challenges new items appeared. Users did not have to do everything at once. They could return at any moment and proceed with the game.

During the game users could earn badges that demonstrate a skill, achievement, or quality. If users successfully answered questions, read information, performed actions, or mastered challenges they received a badge for that. As described we made use of Mozilla's Open Badge Infrastructure for the issuing of badges. So whenever users received a badge a browser window opened, they had to sign in, and then accept the badge. When they did that, the badge was stored in their badge backpack.

5.3 Results and Discussion

At the end of the game the participants were asked to fill in a short questionnaire. In total 12 participants (N=12) completed the questionnaire and thus provided qualitative feedback on the game. Again most questions consisted of 7-point Likert-type scale items giving choices ranging from 1 (not at all) 7 (completely). As expected the results show that participants are highly concerned about the amount of energy they are using at the workplace (M = 5.42), especially regarding the environmental costs, such as higher environmental pollution. They are also highly concerned with what they can do personally to reduce their energy consumption at the workplace (M = 5.75) and performed the suggested energy saving tips. When asked why they are not doing more to reduce their energy consumption at their workplace the participants opted again for more information and detailed feedback on their personal consumption. Compared to the second iteration less participants stated that they need more incentives to save energy thus emphasising the influence of the gamification as incentive mechanism. The majority of participants is highly motivated to take more actions to further reduce their energy consumption at the workplace (M = 5.08).

When asked to evaluate the game the participants stated that the gamification was appealing (M = 4.92). Overall the participants liked “active” game elements, such as action, challenge, and activity most. The “informational” elements, such as information and video were less popular, while badges ranged in between the two. Regarding the expected behaviour change, participants stated that the game in general changed their energy consumption behaviour (M = 4.25), while the information and the activity elements were assigned with the highest potential to do so. Regarding the environmental consciousness, participants stated that the game enhanced their environmental consciousness ($M = 4.67$). In this regard the information and the video element were assigned with the highest potential to do so. Participants stated that the "active" game elements had a slighter higher potential to change energy consumption behaviour compared to the "informational" elements and vice versa for enhancing the environmental consciousness. The badge and the prizes element were in general assigned with the lowest potential, while the potential to change the consumption behaviour was higher compared to the potential to enhance environmental consciousness. All results depicting the potentials are compiled in Table 2. Overall the third iteration showed us the possible impact of gamification mechanisms and the role of incentives.

Table 2. Rated potential of single game elements to increase energy consumption behaviour and environmental consciousness

Game Element	Energy Consumption Behaviour (Mean)	Environmental Consciousness (Mean)
Information	5.50	5.67
Video	4.42	4.83
Action	4.33	4.08
Challenge	4.33	4.17
Activity	4.58	4.42
Badge	3.92	3.42
Prizes	3.17	2.83

6 Results and Discussion Across Iterations, Limitations and Future Research

The presented paper reports about pervasive interventions to increase pro-environmental awareness, consciousness, and learning at the workplace presented within a design-based research cycle with three iterations. The first iteration introduced public displays to increase the awareness for pro-environmental behaviour and energy saving potential. The results revealed the influence on awareness, confidence, and knowledge, but also asked for more personalised and direct feedback. Consequently the second iteration fostered personalised feedback about individual energy consumption at the workplace using different means. On the one hand the results showed the effectiveness and revealed the favoured kind of feedback, on the other hand participants asked again for more information and instructions to initiate conservation activities combined with the need for more incentives to sustain this behaviour. Finally the third iteration focused more on behavioural approaches combined with a motivational and social influence approach utilising gamification and clear incentives. The results underpinned the role and impact of these mechanisms. To sum it up, results of all iterations have provided information on different levels: For the organization the pilots have provided a good guideline how effective energy conservation at the workplace can be enabled and rewarded for employees. For our research we could collect feedback about important design decisions that will influence a large scale pilot, combining the most promising components of the single iterations, i.e. public displays to distribute information, individual displays with personalised feedback, gamification to sustain behaviour change, clear incentives and active game elements, etc. Again this pilot will have to cope with the same limitations as the single iterations. Especially the longitudinal effects of the pervasive interventions need to be examined in long-term studies. Conclusively the presented and future pervasive interventions to increase pro-environmental awareness, consciousness, and learning at the workplace should be considered as continuous campaign rather than one time only activity. Thus companies and organisations should be clear on where to set the focus of the campaign, on whom to address, and on how to balance cost and effectiveness.

Acknowledgement. The presented projects have been partially funded by a SURFnet innovation grant for sustainable ICT solutions and partially by the Centre for Learning Sciences and Technologies (CELSTEC) of the Open Universiteit in the Netherlands.

References

1. Siero, F., Bakker, A., Dekker, G., van den Burg, M.: Changing Organizational Energy Consumption Behaviour Through Comparative Feedback. Journal of Environmental Psychology 16(3), 235–246 (1996)
2. Lo, S.H., Peters, G.-J., Kok, G.: A Review of Determinants of and Interventions for Proenvironmental Behaviors in Organizations. Journal of Applied Social Psychology 42(12), 1–35 (2012)
3. Kollmuss, A., Agyeman, J.: Mind the Gap: Why Do People Act Environmentally and What Are the Barriers to Pro-environmental Behavior? Environmental Education Research 8(3), 239–260 (2012)
4. Börner, D., Storm, J., Kalz, M., Specht, M.: Energy Awareness Displays: Motivating Conservation at the Workplace Through Feedback. International Journal of Mobile Learning and Organisation 6(3/4), 189–203 (2012)
5. Foster, D., Lawson, S., Wardman, J., Blythe, M., Linehan, C.: Watts in It for Me?: Design Implications for Implementing Effective Energy Interventions in Organisations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2357–2366. ACM Press, New York (2012)
6. Cobb, P., Confrey, J., di Sessa, A., Lehrer, R., Schauble, L.: Design Experiments in Educational Research. Educational Researcher 32(1), 9–13 (2003)
7. Reeves, T.C.: Design research from the technology perspective. In: Akker, J.V., Gravemeijer, K., McKenney, S., Nieveen, N. (eds.) Educational Design Research, pp. 86–109. Routledge, London (2006)
8. Design-Based Research Collective: Design-Based Research: An Emerging Paradigm for Educational Inquiry. Educational Researcher 32(1), 5–8 (2003)
9. Börner, D., Kalz, M., Specht, M.: Thinking Outside the Box – A Vision on Ambient Learning Displays. International Journal of Technology Enhanced Learning 3(6), 627–642 (2011)
10. Börner, D., Kalz, M., Specht, M.: Beyond the Channel: A Literature Review on Ambient Displays for Learning. Computers & Education 60(1), 426–435 (2013)

11. Pousman, Z., Stasko, J.: A Taxonomy of Ambient Information Systems: Four Patterns of Design. In: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 67–74. ACM Press, New York (2006)
12. Shen, X., Vande Moere, A., Eades, P., Hong, S.: The Long-term Evaluation of Fisherman in a Partial-attention Environment. In: Proceedings of the 2008 Conference on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization, pp. 1–6. ACM Press, New York (2008)
13. Mory, E.H.: Feedback research revisited. In: Jonassen, D.H. (ed.) Handbook of Research on Educational Communications, pp. 745–784. Lawrence Erlbaum Associates Publishers, Mahwah (2004)
14. Holmes, T.G.: Eco-visualization. In: Proceedings of the 6th ACM SIGCHI Conference on Creativity & Cognition, pp. 153–162. ACM Press, New York (2007)
15. Vining, J., Ebreo, A.: Emerging theoretical and methodological perspectives on conservation behaviour. In: Bechtel, R.B., Churchman, A. (eds.) Handbook of Environmental Psychology, pp. 541–558. Wiley, New York (2002)
16. Werbach, K., Hunter, D.: For The Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press, Philadelphia (2012)
17. ARLearn, `http://code.google.com/p/arlearn/`
18. Mozilla Open Badges, `http://openbadges.org`

3D Interactive Applications on Tablets for Preschoolers: Exploring the Human Skeleton and the Senses

Antonia Cascales[1], Isabel Laguna[2], David Pérez-López[3], Pascual Perona[3], and Manuel Contero[3]

[1] Universidad de Murcia, Avda. Teniente Flomesta 5, 30003 Murcia, Spain
antonia.cascales@um.es
[2] Universidad de Alicante, Cra. San Vicente del Raspeig s/n, 03690 San Vicente del Raspeig, Spain
isabel.laguna@ua.es
[3] Instituto de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano (I3BH), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
{dapelo,pperona,mcontero}@i3bh.es

Abstract. Early years education is an important aspect for the future success of children in the education system. From this perspective, this paper describes the results of a study with preschool children using an interactive learning application on tablets. The project is arranged according to a three-phase process to promote the development of: (1) emergent literacy, (2) digital access for early years learners and (3) basic concepts in knowledge of the environment. The study was conducted with six classes of 87 students aged between 3 years to 6 years, over a 6-week period. During this period, the students were introduced to and engaged in the knowledge of the human skeleton and five senses by using a 3D interactive application on tablets. The quasi-experimental design was based on a nonequivalent groups pretest and posttest design. The interactive learning application was designed around three distinct interaction modes: presentation, exploration, manipulation and evaluation. These phases provided scaffolding for the students to engage with the technology and for the class teacher to develop her own skills. The results on the normalization tests for both control/experimental groups before the experiment were similar. The results after the experiment indicate that students who worked with tablets showed a slight improvement in results of learning outcomes.

Keywords: interactive 3D application, tablets, preschool, teaching/learning process.

1 Introduction

Early educational intervention promotes child advance and school success. The OECD [1] reports that students that attended preschool for one year or more score significantly more in the PISA test than students who did not do. It is important to have good early learning experiences to prepare compulsory education and foundation

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 71–83, 2013.

for life-long learning. Young children who are involved in positive and stimulating experiences and relationships with other children and environment are well equipped to reach their potential in life. In this respect, according to [2], [3], and [4], interaction with technology makes students participate actively in the learning process, promotes better understanding of the instruction, and improves the learning outcomes.

In this context, this work tries to contribute to explore the educational benefits of interactive 3D graphical applications in tablets. This technology offers many potential benefits, especially to motivate young students that show a very positive attitude to this kind of device and give the opportunity to share learning among peers. In this sense, tablets seem a good learning support where students can be provided with interesting environments to learn [5] and encourage collaborative work in the classroom. With this background, this work tries to contribute to the improvement of early childhood education by means of the following objectives:

- Promote educational innovation by a gradual change in teaching methodology in order to utilize the advantages provided by natural interaction with 2D and 3D contents.
- Analyze the possibilities that tablet applications can have on early childhood education to improve interaction between peers.
- Restructure the classroom environment to incorporate tablets devices in daily activities with a constructivist approach to learning.
- Assess improvements that can promote the learning of the students.

Therefore, we have tried to answer the following questions:

- What happens in the teaching and learning process when 3D interactive applications on tablets are used in the classroom?
- How can tablet educational applications help us to achieve the educational objectives?
- What kinds of interactions are produced when this technology is implemented in the classroom?

In order to answer the previous questions we have developed some educational contents and implemented a teaching/learning strategy around them that has been tested on a real preschool scenario using tablet devices.

1.1 Tablets in Preschool and Kindergarten

In recent years, interest in a more natural approach to interact with computers has gained momentum. Thus, progressively the traditional WIMP paradigm is evolving to a more natural interface, such as multitouch interaction. Some of these interface innovations come from new devices such as tablets and tabletops systems.

Tablets are devices, which allow portability and promote cooperation and collaboration though sharing activities that are very interesting from an educational point of view [6, 7, 8].

According to Kearney [9], a technological application is not only for fun. Tablet educational applications should been designed to include aspects that are relevant to the child's development: social experiences, expressive tools and control; so they can help children in their motor-skill and cognitive development. Nevertheless, it is important not to forget that entertainment and fun enhance children's interest and learning.

Couse and Chen [10] studied the viability of tablet computers in early education by analyzing preschool children's ease in adapting to tablet technology and its effectiveness in engaging them to draw. The study found significant differences in level of tablet use between sessions, and engagement increased with age. Participant teachers stated high child interest and children quickly developed ease with the stylus for drawing. Rankothge et al. [11] conducted a study on the introduction of a technology assisted tool for the learning skills development in early childhood. The final outcome was a Tablet PC based application to help the children in their learning experience at early ages. The developed tool improved the writing and speaking skills of the participant children in an entertainment based way.

Sandvik et al. [12] concluded that tablets devices were able to raise kindergarten children language and literacy skills through interaction with an image repository. It was tested that children developed the ability to pick up elements from the real-world contexts and connect them to technology.

A common trend in the previous works is tablets promote that students share and help one another, ask for and provide information and explanations, and collaborate to solve problems. However, we know relatively little about how the use of interactive 3D contents on tablets can enhance and promote peer learning for preschoolers and help to create a constructivist learning environment. We interpret learning in early years as Papert [13] emphasizing active construction of knowledge and understanding. In this context, learning is seen as an active process of knowledge construction. Constructivist learning theory focuses on learning process instead of the content. For that reason, it emphasizes active knowledge through meaningful activities, interaction and communication with peers. This pedagogical model gives a more active role for students in their own learning process. 3D interactive applications on tablets can provide learning environments that support meaningful learning activities and interaction between peers.

In the next sections, a detailed description of the didactic contents is provided. Then, the experimental design is presented, followed by the results, discussion and conclusions.

2 Materials and Methods

2.1 Didactic Materials

In this study 10 inches low cost Android tablets were selected as the hardware platform. Performance and cost were the main reasons to select this equipment. The didactic application consisted in a content launcher, and an installation tool. It was also provided an installation manual to promote its easy set up.

The installation tool consists of a conventional apk file, which can be downloaded from Android's Play Store. Nevertheless, the installation was expected to be set up by a teacher, who launched the selected content for his students.

The application structure consists of three different parts: a launcher menu, which is designed to manage a collection of 2D and 3D interactive contents. It also allows the user to launch a specific educational module by selecting a specific content and language through different menus. The launcher was originally designed to be used by any user; however, in this study only teachers managed it. In fact, it could be used directly by children, but this was considered irrelevant for the finality of this study and for this reason this functionality was limited.

Application's interface allows kindergarten children to navigate through different information. Users are allowed to freely select the information to show, and they are also enabled to return to previous information if they wanted. Besides, just by touching interface elements like buttons and scroll bars, the system enables to control an auditory narrative (play, pause and stop), provides a user help guide, and a context-sensitive help on the action that is being performed at the moment. Moreover, the interface allows the visualization of models and procures interaction with 2D and 3D models. Furthermore, GUI information can be hided or shown to improve 3D model explorations by dragging labels in opposite directions. Therefore, when every element of the interface is shown, left side buttons can give access to different activities, as presented below (Fig.1). In addition, in the lower side of the screen, there is a rectangle reserved to show detailed explanations in text format. Note that those explanations are also shown as an auditory narrative. Also, note that this text box incorporates an automatic scroll to support long texts.

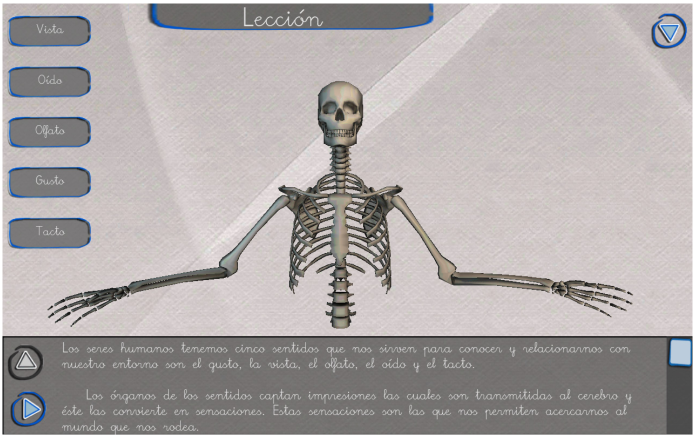


Fig. 1. 3D Interactive Application where some of its interface elements are shown deployed

The last part is the main view. Main view occupies most of the screen and it shows 3D models, which can be interacted by users directly by touching the screen. 3D models can be dragged, rotated and escalated through natural gestures, enhancing user visualization of models.

The Skeleton Module provides two kinds of activities, "Lesson" and "Exercises". On the one hand, "Lesson" allows user to observe the scene; there is no more interaction than exploring 3D models from different points of view while listening to the corresponding audio. This module is composed of five subactivities, which can be accessed by a set of buttons located at the left of the screen: “Bones system”, “Bone Joints”, “Skeleton” and “Types of Bones”. Each sub-activity presents the most relevant information related to its topic. In order to control the user interaction, user navigates through menus pressing buttons (Fig. 2).

If the user would want to return to previous information, he just had to press the escape button of tablet device. On the other hand, “Exercises” are a compilation of five games: “Joints classification”, where user has to classify the different bone joints between its kinds; “Touch the Joint”, in this game students are asked to touch a specific joint while a skeleton is moving; “Fix the skeleton”, in this puzzle-like game children must fix the skeleton dragging its parts to the proper place; “Bones types”, in this game it is shown different glowing bones and students must classify them; and “Touch the Bone”, where a complete skeleton is shown, and students are asked to touch the correct bone. In every game, it is also shown information about correct and wrong answers to boost children self-learning and autonomy.

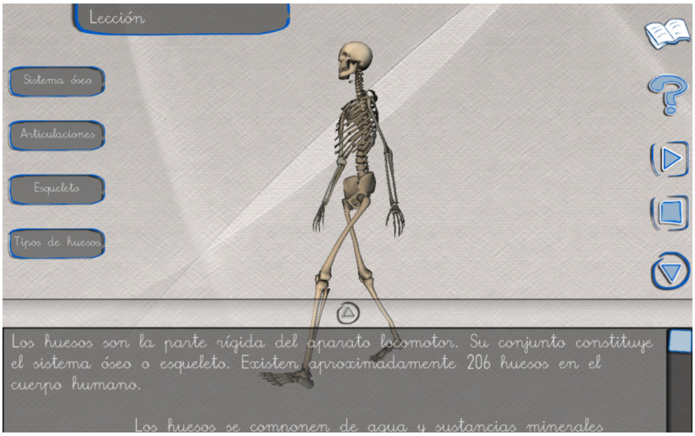


Fig. 2. Skeleton Module. GUI elements deployed.

In a similar way than in the Skeleton Module, the Senses Module provides two kinds of activities, “Lesson” and “Exercises”. In this case, in Lesson mode the contents taught are the five senses: “sight” (Fig. 3), “taste”, “hearing”, “smell” and “touch”.

In “Exercises” mode are proposed two games: “Association”, where user has to associate different objects to the most relevant sense which allows its recognition; and “Composition”, where students are asked to match a sense with different parts of a skull. It is important to mention that the two applications resources were initially designed for primary school students, but after a preliminary evaluation with preschoolers, it was observed that they were able to access the basic functionality of the application without any problem.

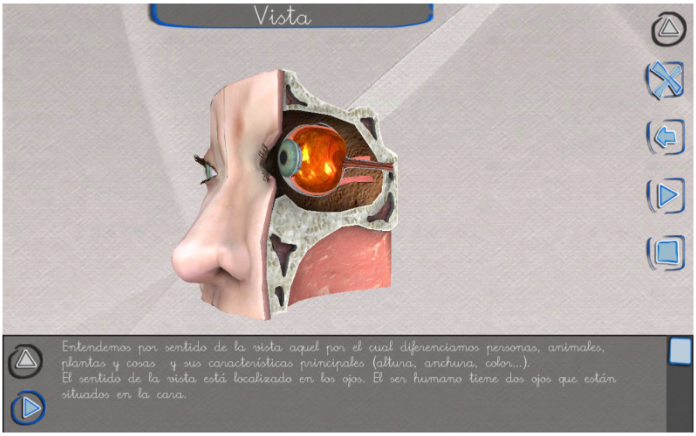


Fig. 3. Interface characteristics. Senses Module.

2.2 Participants

The research involved six groups of eighty seven preschoolers, with ages between three and five years from the public school Virgen de los Desamparados in Orihuela (Spain) and their teachers. All groups belonged to the second cycle of pre-primary education, according to the Spanish education system, but they were from different level. The sample consisted of: two groups of three-year-olds, with 24 students; two groups of four-year-olds, with 30 students; and two groups of five-year-olds, with 33 students (Table 1). One group of each age was taken as the control group, while the other was taken as the experimental group. The participant teachers had received the same training.

Table 1. Demographic Information of Subjects by Age Group

	Control Group			Experimental Group		
		Gender			Gender	
Age Group	**N**	M	F	**N**	M	F
3 to 4 years old	**12**	7	5	**12**	6	6
4 to 5 years old	**15**	8	7	**15**	7	8
5 to 6 years old	**16**	11	5	**17**	10	7
Total	**43**	26	17	**44**	23	21

The school is located in a rural area. It is one of the seventeen pilot technological schools in the province of Alicante (Spain). The school is fully equipped with technology and count on a team of teachers experienced in ICT, which works hard to improve the use of ICT in school. Regarding the students participating in our research, they have been using ICT in school since they came to it; therefore, there are children who have been using ICT in class previously and others which are having their first experience with the ICT in this course. Students work every day with technology in their own classrooms, where they have several computers and an interactive whiteboard.

2.3 Experimental Design and Method

In this research, a nonequivalent group pretest and posttest [14] design has been chosen. Under this scheme, one group (the experimental group) received the intervention consisting in using the 3D interactive applications on tablets (students from the experimental group can be seen in Fig. 4), while the other group (the control group) does not use it. The intervention was done in a natural situation, without a random selection of groups [15]. Initial conditions for all groups weren´t similar: each group was composed by a different number of children. In addition, its relationship with the ICT was very different. None of the groups had studied the topic of the Human Skeleton and the Senses previously.

This experience has been developed using an active methodology, based on communication and research. This approach entails taking an area of interest –the Human Skeleton and the Senses– and using these topics as a basis for an in-depth enquiry or research. Areas of learning are not simply linked by a topic; they are integrated as a result of the investigative process. Knowledge and skills are not taught in isolation, but rather acquired and practiced within a meaningful context that makes sense to children. The methodology of work is representative of a pedagogy that stems from a positive image of the young child as a competent learner who is capable of taking an active role in their education. The provocation shows the inclusive potential and how they can enthuse children of all aptitudes and abilities, as well as motivate children to want to learn by building upon their interests.

On having used this methodology, the teachers were deeply implied providing feedback data about student experiences, bearing in mind the age and evolutionary characteristics of the pupils. The chosen didactic unit for all groups was “the Human Skeleton and the Senses”. Two versions of these didactic materials were created. The only difference between them was that the “experimental unit” provided the 3D interactive applications described in the previous section. In this way, both units have the same educational curriculum content. Therefore, the independent variable of this research was the presence of 3D interactive applications on tablets as a didactic tool, and the dependent variable was the ease of use.

On the other hand, all preschool students worked properly following the didactic guides developed by participant teachers. However it was the first time they used the tablets in the classroom. The experimental groups were divided into three groups and the participant teachers familiarized each group with the 3D interactive application on tablets. The participant teachers demonstrated how to use the application of sense (lesson and exercises) and each child spent a few minutes practicing with it on the tablet. The working sessions with tablets had a duration of three weeks and students worked in pairs and small groups (4 or 5 children).

Fig. 4. Children using 3D interactive applications on tablets

The assessment of experimental and control groups was performed using an evaluative categorical scale completed by teachers (see Table 2). This scale consisted of 20 items, where each item was checked according to the following categories: A (Achieved), IP (In Progress) and NA (Not Achieved). This test was performed twice, before implementing the program (pretest) and after implementing the program (post-test).

Table 2. Students' categorical estimation scale

Pupil:	NA- Not Achieved IP- In progress A- Achieved		
Item		**Criteria**	
Skeleton part			
• Child recognizes backbone	NA	IP	A
• Child recognizes cranium	NA	IP	A
• Child recognizes mandible	NA	IP	A
• Child recognizes scapula	NA	IP	A
• Child recognizes clavicle	NA	IP	A
• Child recognizes ribs	NA	IP	A
• Child recognizes humerus	NA	IP	A
• Child recognizes radius	NA	IP	A
• Child recognizes ulna	NA	IP	A
• Child recognizes femur	NA	IP	A
• Child recognizes tibia	NA	IP	A
• Child recognizes fibula	NA	IP	A
Knowledge about joints:			
• Child recognizes wrist	NA	IP	A
• Child recognizes elbow	NA	IP	A
• Child recognizes fingers	NA	IP	A
• Child recognizes knee	NA	IP	A
• Child recognizes ankle	NA	IP	A
• Child recognizes heel	NA	IP	A
Knowledge about sight:			
• Child recognizes iris	NA	IP	A
• Child recognizes pupil	NA	IP	A
• Child recognizes retina	NA	IP	A
• Child recognizes optic nerve	NA	IP	A

Table 2. *(continued)*

Knowledge about hearing:			
• Child recognizes outer ear	NA	IP	A
• Child recognizes middle ear	NA	IP	A
• Child recognizes inner ear	NA	IP	A
• Child recognizes cochlea	NA	IP	A
Knowledge about taste:			
• Child recognizes salty	NA	IP	A
• Child recognizes sweet	NA	IP	A
• Child recognizes bitter	NA	IP	A
• Child recognizes sour	NA	IP	A
Knowledge about smell:			
• Child recognizes nasal cavity	NA	IP	A
• Child recognizes olfactory bulb	NA	IP	A
• Child recognizes cartilage	NA	IP	A
• Child recognizes nose	NA	IP	A
Knowledge about touch:			
• Child recognizes epidermis	NA	IP	A
• Child recognizes dermis	NA	IP	A
• Child recognizes sweat gland	NA	IP	A
• Child recognizes nerve	NA	IP	A

Besides, there was an assessment of levels of tablet use. It was also performed using an evaluative categorical scale and it was completed by teachers too. Levels of tablet use were coded according to Table 3.

Table 3. Definitions levels of Tablet Use

	Definition
Level 1: Explore/Experiment	Child tries to figure out what the tablet can do, clicking on different options to see what will happen if….
Level 2: Investigate	Child tries to figure out how to use the tablet to obtain information (e.g., How can I know the names of the bones?)
Level 3: Apply the knowledge	The child puts in function the knowledge acquired to realize the tasks.

Finally, there was an interview with participant teachers about how 3D interactive applications helped to create constructivist learning environments. In this interview, some questions were related to interaction capabilities of the developed 3D applications and others related to the student learning process.

3 Results

Total scores for pretest and posttest in both experimental and control group are presented in Table 4.

Table 4. Total scores for pretest and posttest (absolute frequencies)

	Control Group			Experimental Group		
3 to 4 years old	A	IP	NA	A	IP	NA
TOTAL Pretest	2	26	56	2	27	55
TOTAL Postest	17	32	35	31	41	12

	Control Group			Experimental Group		
4 to 5 years old	A	IP	NA	A	IP	NA
TOTAL Pretest	9	19	72	10	17	73
TOTAL Postest	24	34	42	38	48	14

	Control Group			Experimental Group		
5 to 6 years old	A	IP	NA	A	IP	NA
TOTAL Pretest	14	20	78	14	22	76
TOTAL Postest	27	42	43	49	45	18

Table 5 summarizes the results obtained from applying the categorical estimation scale about levels of tablet use of in the experimental groups as defined in Table 3.

Table 5. Levels of use Tablet in unit "Skeleton and the Senses" (absolute frequencies)

	Experimental Group		
	1	2	3
3 to 4 years old	4	4	4
4 to 5 years old	4	5	6
5 to 6 years old	4	6	7
Total	12	15	17

Table 6 presents the detailed posttest results obtained from applying the categorical estimation scale presented in Table 2. Experimental group reflects a slight improvement with respect to control group although there is no statistical significant difference.

With respect to opinions expressed by participant teachers, it is important to note teachers provided evidences of children's interest and the feasibility of 3D applications on tablets for preschooler to create constructivist learning environments. All teachers perceived more interest in children and an improved interaction between peers and small groups. In general, all they agreed that tablets promotes dialogue and communication in learning situations, and students required less attention from teacher, as they worked in a more autonomous way. Regarding to the learning process, participating teachers in this study observed that 3D interactive apps helped to build a more meaningfully learning, where pupils learn by doing and discovering by themselves. Students were engaged in the construction of their own learning, and therefore the preschoolers developed a range of skills that will serve for future daily life. Other comments from participant teachers noted the excitement of children when they knew that they were going to use the 3D apps with the tablets and their curiosity to learn exploring these 3D environments.

Table 6. Learning outcomes in didactic unit "Skeleton and the Senses" (absolute frequencies)

	Control Group			Experimental Group		
3 to 4 years old	A	IP	NA	A	IP	NA
Skeleton part	2	5	5	4	7	1
Joints	2	4	6	5	5	2
Sight	3	5	4	5	6	1
Hearing	3	5	4	4	6	2
Taste	2	5	5	4	5	3
Smell	2	4	6	4	6	2
Touch	3	4	5	5	6	1
	Control Group			Experimental Group		
4 to 5 years old	A	IP	NA	A	IP	NA
Skeleton part	3	6	6	5	8	2
Joints	3	5	7	6	6	3
Sight	4	6	5	6	7	2
Hearing	4	6	5	5	7	3
Taste	3	6	6	5	6	4
Smell	3	5	7	5	7	3
Touch	4	5	6	6	7	2
	Control Group			Experimental Group		
5 to 6 years old	A	IP	NA	A	IP	NA
Skeleton part	4	6	6	7	8	2
Joints	3	6	7	8	6	3
Sight	4	6	6	6	9	2
Hearing	5	6	5	7	7	3
Taste	3	7	6	6	9	3
Smell	4	5	7	7	7	3
Touch	4	6	6	8	7	2

4 Discussion and Conclusions

Young children of 3 to 4 year old were able to quickly learn to use the 3D interactive applications on tablets as a medium for representing their ideas and to improve their learning. The children in this study were able to become comfortable using the 3D applications when adults had given them some instructions and when they worked collaboratively with their peers. As the children gained familiarity with the tablet, they became more independent, asking for less instruction and assistance from adults. This was to be expected, as independence leads to a deeper exploration and a fuller utilization of the technology to represent ideas in a more productive way. As a result, it has been made easier to encounter more situations that are new. Finally, the use of ICT did not influence the ease with which children became acclimated to this new technology.

The slight improvement in results of learning outcomes in didactic unit "Skeleton and the Senses" must be put in context. 3D interactive applications on tablets were one of the components of the didactic activities designed to support this learning unit. The learning process was organized around team work, and the 3D interactive applications were used by the preschoolers at their own pace and depending on their evolutionary characteristics. Contents in tablets served as a catalyst providing a real motivation and stimulus for the children, and teachers observed a very positive impact on students.

Participant teachers had no previous exposure to this technology, but it was easily integrated in the class dynamics. The perception of participant teachers was that 3D interactive applications on tablets improved considerably learning activities and supported students' constructive learning approach. Teachers considered the 3D applications a resource tailored to the characteristics of their students and thus it was useful for learning. When teachers provided social facilitation for children using the tablets in the form of scaffolding and scripting the environment, positive peer interaction significantly increased. The children reflected upon the interactive content before sharing and discuss about it with each other.

All study participants considered that the use of 3D interactive applications on tablets is a good tool in the teaching-learning process. The conclusions we reached in our experience of inclusion of tablets as a part of the teaching-learning process are the following:

- The use of tablets promotes opportunities to investigate different interests of students and for developing skills in applying knowledge to informal reasoning. The work of teaching improves with the use of these devices. Daily work is more playful and fun for both students and teachers.
- In spite of the age of the students, the preschool children learn more when they are using the 3D interactive applications on tablets and they also achieve more learning goals than if they are not using.
- 3D interactive applications on tablets promote communication skills. In this sense, all kind of interactions in the classroom are promoted: between teacher and students; students and students; students and families; families and families; and teachers and teachers.

The experience with the use of tablets has been very constructive for teachers and students. After this experience, participant teachers will encourage their colleagues to use it in their classrooms. With regard to the pupils, the engagement with technology does not appear to be a simple function of age, but rather a more complex relationship between technology characteristics and the use done of it.

Acknowledgements. The Spanish Ministry Economy and Competitiveness partially supported this work (Project ref. TIN2010-21296-C02-01).

References

1. OECD: Investing in high-quality early childhood education and care (ECEC), http://www.oecd.org/dataoecd/0/28/48980282.pdf
2. Chou, C.: Interactivity and interactive functions in web-based learning systems: A technical framework for designers. British Journal of Educational Technology 34, 265–279 (2003)

3. Crowther, K., Waddoups, G.L.: Improving the quality and effectiveness of computer-mediated instruction through usability evaluations. British Journal of Educational Technology 35, 289–303 (2004)
4. Hinostroza, J.E., Mellar, H.: Pedagogy embedded in educational software design: Report of a case study. Computers & Education 37, 27–40 (2001)
5. Kearney, J.: Educating Young Children - Learning and teaching in the early childhood years. Early Childhood Teachers' Association (ECTA Inc.) 3(18) (2012)
6. Marco, J., Cerezo, E.: Bringing tabletop technologies to kindergarten children. In: HCI 2009 International Conference on Human-Computer Interaction– Celebrating people and technology, pp. 103–111. Springer, Heidelberg (2009)
7. Heft, T.M., Swaminathan, S.: Using computers in early childhood classrooms: Teachers' attitudes, skills and practices. Journal of Early Childhood Research 6(4), 169–188 (2006)
8. Wang, X.C., Ching, C.C.: Social construction of computer experience in a first-grade classroom: Social processes and mediating artifacts. Early Education and Development 14(3), 335–361 (2003)
9. Kearney, J.: Educating young children: Learning and teaching in the early childhood years. Early Childhood Teachers' Association (ECTA Inc.) 3(17), 35–38 (2012)
10. Couse, L.J., Chen, D.W.: A Tablet Computer for Young Children? Exploring Its Viability for Early Childhood Education. Journal of Research on Technology in Education 43(1), 75–98 (2012)
11. Rankothge, W.H., Sendanayake, S.V., Sudarshana, R.G.P., Balasooriya, B.G.G.H., Alahapperuma, D.R., Mallawarachchi, Y.: Technology assisted tool for learning skills development in early childhood. In: 2012 International Conference on Advances in ICT for Emerging Regions (ICTer), pp. 165–168. IEEE (2012)
12. Sandvik, M., Smørdal, O., Østerud, S.: Exploring iPads in practitioners' repertoires for language learning and literacy practices in kindergarten. Nordic Journal of Digital Literacy 3(7), 204–221 (2012)
13. Papert, S.: A critique of technocentrism in thinking about the shool of the future. MIT.Media, Cambridge (1990)
14. Cook, T.D., Campbell, D.T., Day, A.: Quasi-experimentation: Design and Analysis Issues for Field Settings, pp. 19–21. Houghton Mifflin, Boston (1979)
15. Buendía, L., Colás, P., Hernández-Pina, F.: Métodos de Investigación en Psicopedagogía. McGraw Hill, Madrid (1997)

NLP-Based Heuristics for Assessing Participants in CSCL Chats

Costin-Gabriel Chiru[1], Traian Rebedea[1] and Stefan Trausan-Matu[1,2]

[1] University Politehnica of Bucharest, Department of Computer Science and Engineering, 313 Splaiul Independetei, Bucharest, Romania
[2] Research Institute for Artificial Intelligence of the Romanian Academy, 13 Calea 13 Septembrie, Bucharest, Romania
{costin.chiru,traian.rebedea,stefan.trausan}@cs.pub.ro

Abstract. In this paper, we present an application that can be used for assessing the participants' contributions to multiple chat conversations that debate the same topics according to different criteria (involvement, knowledge and innovation), along with the ranking of the conversations considering a list of important concepts to be debated. As several factors have been used for determining each participant's score, we needed to determine their quality for providing an answer that correlates well with the judgment of human evaluators for the same conversations. Thus, we also propose a methodology for testing the values of different factors that may be used for assessing participants in collaborative conversations in order to identify which of them are better or worse suited for providing automatic assessment. Our analysis showed that the heuristics used to assess participants' innovation and involvement were the best correlated with the human judgment, while at the opposite end was the heuristic used for assessing participants' knowledge.

Keywords: CSCL, NLP, Participant Assessment, Heuristic Testing, Chat Conversations.

1 Introduction

Nowadays, instant messaging (or chat) is one of the most popular ways of collaboratively exchanging information over the Internet in a synchronous manner between a variable number of persons. The importance of chat in people's everyday life is also highlighted by the decision of very large companies (Skype, Yahoo, Google, Facebook, Tinychat, etc.) to built dedicated systems for this kind of communication. Moreover, chat is present in various flavors and designs, such as Internet Relay Chat (IRC), and is even present in social networks and different popular learning management systems such as Moodle, Blackboard, etc.

Chat is also one of the favorite environments for Computer Supported Collaborative Learning (CSCL) tasks that require online and synchronous textual interactions among participants [10, 11]. It has been used for debating, reaching agreements or solving complex problems usually in small collaborative teams. Due to the wide

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 84–96, 2013.

adoption in CSCL activities, chat technologies have been extended with functionalities required by these specific tasks such as the explicit referencing mechanism and the whiteboard facility that are present in Concert Chat [6]. Other CSCL environments have been designed to solve specific problems, but they include an instant messaging module as an important part of the system – for example the chat boxes in the C-CHENE system [1].

In spite of its popularity and of the huge quantities of data that are exchanged daily through chat conversations, most of the existing tools for supporting chats are only aiming at facilitating the conversation without offering analysis instruments. The same is true for CSCL tasks: although chat is widely used for interactions in learning tasks, there are very few systems that were built for analyzing the content of chat conversations used by small groups.

One example is PolyCAFe [7], a system which analyzes each user contribution and provides abstraction, visualization and feedback services for supporting both learners and tutors. The system is built using several Natural Language Processing techniques for identifying implicit links between utterances based on semantic coherence, speech acts and adjacency pairs, and for determining the evolution of the most important topics under discussion using TF-IDF and Latent Semantic Analysis (LSA) techniques. Using the links found between utterances, a conversation graph is constructed that can be used to discover conversation threads and to assess the individual involvement of the participants. Another example is presented in [3] where the authors start from the idea of rhythmicity of topics and of participants' utterances in order to evaluate the quality of the chat and the personal involvement of the participants.

Still, what we consider to be missing from the two presented systems is the fact that although the conversations debate the same topics, they are not compared in any ways. Therefore, there is no easy way to rank the chats conversations or their participants according to some criteria. This is what we want to address in this paper. For this task, we started from the heuristics proposed in [3] and developed two different evaluation methods – an intrinsic one considering only one conversation at a moment, and a multi-chat evaluation, that considers the conversation in the context of the whole corpus from which it has been extracted.

The paper continues with a brief presentation of the theoretical ideas that constitute the basis for the study. Next, we introduce the application that has been developed and the analyzed corpus, along with the factors used for assessing the participants and the chat conversations. Afterwards, we present the main results of our investigation, along with a comparative study of these results with previous systems. The paper concludes with our final remarks.

2 Theoretical Background

One of the main advantages of using chat as a communication environment by small teams solving CSCL tasks is that it allows the formation of several parallel conversational floors [9]. The participants to the conversations make use of this feature and engage into conversations with several discussion threads. In most times, only the best (or few of them) discussion thread continues: the one that gets the attention of most participants. This may be one of the most important reasons why chat conversations are often preferred to face-to-face interactions. However, the presence of multiple discussion threads raises another problem when trying to perform an automatic analysis: disentanglement of the chat's discussion threads [4].

Most of the research done in conversations' analysis is limited to a model with two interlocutors where a single topic is in focus any time. In this case, usually the analysis employs rhetorical structure theory or adjacency pairs [5]. However, in order to solve the problem of parallel discussion threads arising in multi-party chat conversation, a new model may be needed for performing a well-tailored discourse analysis. One of these proposed models is based on Bakhtin's ideas [2] – especially inter-animation, the dialogism theory and polyphony in discourse – that allow the understanding of the collaboration mechanisms and provide the means to measure the contributions of participants. Bakhtin [2] considered that in any text there is a co-occurrence of several voices that gives birth to inter-animation and polyphony: "Any true understanding is dialogic in nature." [14]. The same idea is expressed in [13] building on the work of Schegloff [8]: "knowledge is socially built through discourse and is preserved in linguistic artefacts whose meaning is co-constructed within group processes". The idea of voice is extremely important in Bakhtin's theory although he doesn't explicitly mention how a voice could be identified and quantified objectively in a given discourse.

In this paper we investigated the concept repetition in order to capture the information that we believe is the closest to the Bakhtin notion of voice. In doing so, we also considered the importance of repetition expressed by Tannen [12]: "dialogue combines with repetition to create rhythm. Dialogue is liminal between repetitions and images: like repetition is strongly sonorous". However, it is not merely a repetition of words, but rather of lexical chains as we consider that all the words in a lexical chain refer to the same idea (or voice, in a dialogic analysis). Starting from repetition and other qualitative measures for participation in chats, we have proposed a set of heuristics used to assess the contribution of each participant. The main advantage of the current idea over the previous ones is that it allows the assessment of a participant in two different situations: using only the chat one was part of or comparing with all the chats that have been generated by different teams working for the same CSCL task.

3 Application for Automatic Assessment

The main purpose of this application was to propose some criteria for automatically assessing the contribution of the participants in a chat, and afterwards to provide the meanings for comparing the chats one to each other and for ranking all the considered participants according to the considered criteria. In order to do so, we started from the heuristics proposed in [3] and grouped them so that to respond to three different criteria: involvement, knowledge and innovation. We have also built an overall heuristics to combine the three criteria from above and to give a mean to the search for 'the best candidate' from the group.

3.1 Corpus

The application was used to assess chat conversations created by senior year undergraduate students involved in a Human-Computer-Interaction (HCI) class. The students were asked to debate about different web-collaboration technologies (forums, blogs, chats, wikis, google wave, etc.), highlighting the weaknesses and strengths of the existing tools and eventually devising a way to combine these tools in order to obtain an instrument that would be useful for sharing information and collaboration in a company. Each participant had to study individually the given technologies in order to identify their advantages and problems before the conversation started and afterwards they had to choose one of them and to support it in front of the other participants during the chat.

The main purpose of these conversations was to ease the learning process about the considered platforms by providing each participant the possibility to critique the others' platforms and in the same time to defend its own. This way, they were able to 'see' the platforms from different perspectives and had the chance to make a comparative analysis of all the considered platforms. Finally, they had to combine the existing technologies and to develop use case scenarios so that the advantages offered by each of them to be exploited to the maximum and in the same time to eliminate the identified problems related to each individual technology.

The corpus used for validating the results of the system consists of 7 chat conversations ranging from 248 to 524 replies per conversation, for a total of 2514 utterances. Each of the 7 discussions was created by a small group of 5 students, resulting in 35 students who participated to this experiment. The same data has been used in validating previous systems and was collected within the FP7 LTfLL project [7].

3.2 Heuristics Used for Participants' Evaluation

In order to be able to assess the contribution of each participant to the conversation, we started from some of the heuristics that were previously suggested in [3]:

- *Number of replies* indicating how interesting the conversation is for the considered participant – a high value is desired;
- *Activity* of each user showing how elaborated one's replies are. It is expressed as the average number of characters per reply and we considered that the higher this value is, the better is the user's activity;
- *Absence* from the conversation, probably due to the lack of interest for what was debated. This heuristics is computed as the average number of utterances that could be found between two interventions of the user of interest – should be as low as possible;
- *Persistence* of the user in the conversation expressed as the average number of consecutive utterances. Small values are preferred for this heuristic since it expresses the participant's intention of 'keeping the floor' (therefore, monopolizing the conversation) and we aim for an interactive conversation;

- *Repetition* of other participants' concepts offering insights about how much attention a participant devoted to the content uttered by the others – high values are preferred;
- *Usefulness* of the participant in the conversation stating how much the other participants benefited from this user's replies. It is expressed as the percentage of the concepts introduced by the considered user that were further used by the other participants. Higher values are better;
- *On topic* saying how devoted the participant was to keep the conversation on the right track, considering the imposed topics for the discussion. This heuristic is computed as the number of on-topic replies uttered by participant divided by the total number of utterance that he introduced.

From the heuristics proposed in [3], we eliminated the *explicit connections* heuristic from our analysis because we considered that it provides an incomplete/incorrect analysis, since there are a lot of other connections in the conversations that remain implicit and cannot be identified, therefore being impossible to also consider them in the analysis.

On the other hand, we investigated two different heuristics – *participant's knowledge* and *participant's innovation*. We computed *participant's knowledge* as the percentage of the concepts introduced by the participant that were semantically connected with the main ones imposed for debating (chat, blog, forum, wiki). In order to identify the concepts that were semantically connected with the imposed ones, we used lexical chains built on top of the WordNet (http://wordnet.princeton.edu/) lexical database. Since these concepts were very specific to the HCI domain, most of the terms that were semantically connected to them were not also connected in WordNet. Therefore, in order to be able to develop good quality lexical chains we had to develop a taxonomy of concepts related to each of the imposed topics and to use them in order to augment the identification of the concepts that were semantically connected to the four initial concepts imposed for debating. Therefore, we ended up with 76 concepts related to chat, 44 related to blog, 63 terms for forum and 31 for wiki. We have chosen to use ontologies (WordNet), in spite of the fact that previous systems used LSA, because of the much reduced availability of conversation corpora that were required for performing LSA. In the same time, this decision increased the possibility of using our application for different topics than the one that it was designed for (CSCL tools), because WordNet is a general ontology that can be applied for any domain. Of course, if the conversations are very specific for a given domain, the WordNet ontology has to be augmented with a self-constructed taxonomy of terms related to the conversations' topic. Moreover, we consider that lexical chains are able to better capture the idea (or voice, in a dialogic analysis) than LSA, as all the words/concepts in a lexical chain are similar to one another, while this cannot be determined by the plain use of LSA. It might also be interesting to investigate whether lexical chains cannot be built using semantic distances from WordNet combined with the semantic relatedness scores computed by LSA.

The second introduced heuristic, participant's innovation, was computed as the number of concepts introduced in the conversation by each participant and represented the degree of new information introduced by that person.

After computing the values for each of the presented heuristics for every conversation in the corpus, we have identified the highest and lowest values for each factor and scaled the original values in order to obtain values in the [0, 1] interval. For doing this, two possible alternatives were considered: one related to a *single chat* and another considering *all the chats* that were used. The first alternative can be used for evaluating the contribution of each participant in a single conversation, while the second option can be used for evaluating a person considering also the activity of all the other persons involved in different conversations that are from the same corpus (assuming that all the conversations in a corpus have been created for solving the same task, but by different groups).

After that, we combined the first five heuristics: number of replies, activity, absence, persistence and repetition considering that they all characterize the involvement of a participant. Therefore, we computed involvement as the mean of the values obtained for each of the 5 considered heuristics. In order to determine the values corresponding to each heuristic, we normalized the obtained value (for a given participant from a given conversation) according to the minimum – maximum values that were observed (in that conversation for the single chat option or in the whole corpus of conversations for the multi-chat evaluation), also taking in consideration the type of the desired values (either small or high). The final (overall) score for each participant was obtained as an average value of normalized *involvement*, *knowledge* and *innovation*.

3.3 Participant Analysis

The application has two major parts: the first one is related to evaluating chat participants from different points of view, considering only the current conversation or all the conversations from the corpus. The user has the option to choose what criterion he wants to use (involvement, knowledge, innovation or overall) and after that he is presented with the results according to the chosen criterion (see Figure 1 where the chosen criterion is participant's involvement). The results are ordered according to the participants' scores and they are presented in both forms – *intrinsic* (considering only the participants' score from the given conversation) and *multi-chat* (presenting the scores obtained by the participants if the whole corpus of chats was considered instead of the single chat).

Table 1 gives an example of evaluating the participants to one of the conversations used for validation (*chat_116*) considering all 4 heuristics: involvement, knowledge, innovation and overall. In this example, one can see that Mona was very involved and the most innovative person, but in the same time she had a poor score for the knowledge heuristic. This situation is explained by the fact that she introduced in the conversation a lot of new concepts that were related to the topics imposed for debating (therefore the high involvement and innovation scores) but in the same time, she did not develop the introduced concepts, this task being fulfilled by corina (the participant with the highest grade for knowledge and in the same time proving low innovation).

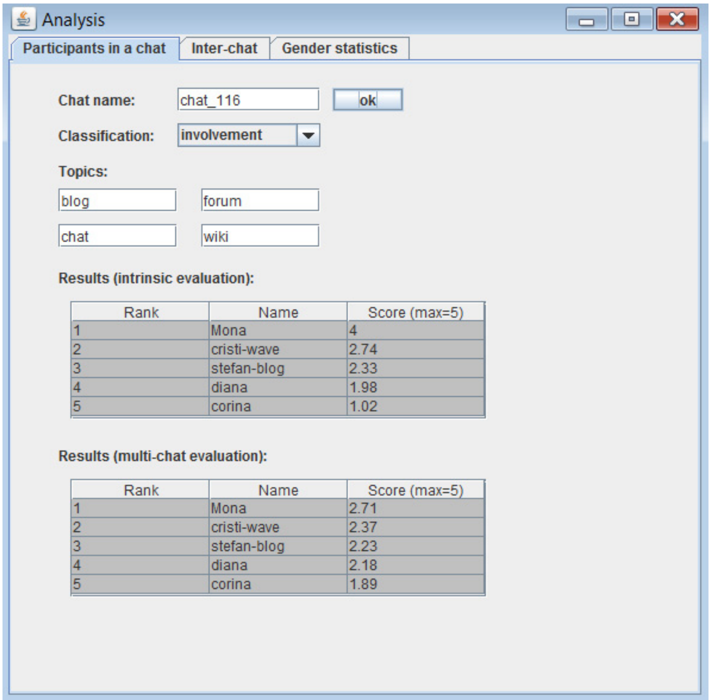


Fig. 1. The results of the analysis considering the criterion “involvement”.

Table 1. Chats length and participants’ distribution

Partici-pants in chat_116	**involvement**		**knowledge**		**innovation**		**overall**	
	single	**multi**	**single**	**multi**	**single**	**multi**	**single**	**multi**
Mona	4	2.71	0.03	0.41	1	0.39	0.61	0.45
cristi-wave	2.74	2.37	0.43	0.63	0.46	0.19	0.48	0.43
stefan-blog	2.33	2.23	0.35	0.58	0.25	0.12	0.36	0.38
diana	1.98	2.18	0	0.39	0.24	0.11	0.21	0.31
corina	1.02	1.89	1	0.94	0	0.02	0.4	0.45

3.4 Chat Analysis

The second part of the application tries to compare the conversations one to each other in order to decide which one of them has achieved the best results considering the fact that the discussion topics were externally imposed by the tutors. Therefore, in this part we investigated the topic rhythmicity (how often each of the four concepts ‘came on the floor’) from each conversation and the vocabulary used in each conversation and in the whole corpus (see Figure 2).

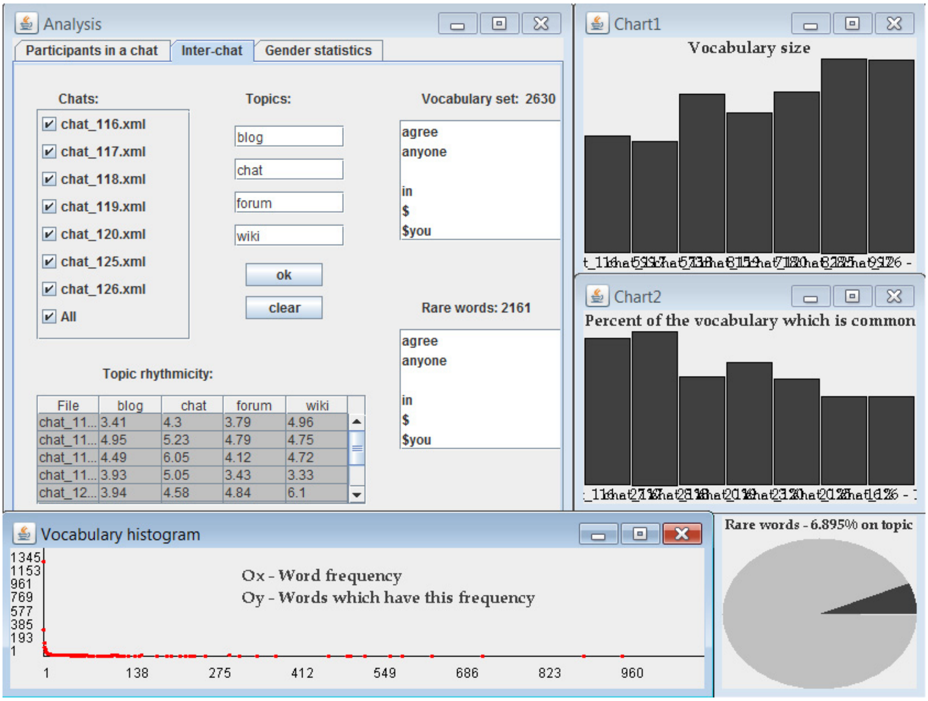


Fig. 2. The results of the analysis considering the criterion "involvement"

The scores for the rhythmicity of the 4 pre-imposed concepts are presented in the left part in a table to be easy to compare the scores (see Figure 2). The graphics present the vocabulary analysis using a frequency histogram: there were 2630 words used in the 7 conversations, out of which 2161 have been used in maximum 2 conversations and therefore were considered to be rare words. Another interesting thing that can be observed is that 6.895% of these rare words (149 words) were considered to be on topic, so a lot of important information was dispersed in only one or two chats. The conversations had 162 words that were found in all the chats, representing between 16 % and 28% of their vocabulary.

4 Analysis and Results

In order to evaluate the quality of the assessment, we asked 4 HCI experts to evaluate the conversations that were part of our corpus [7]. We ended up with 15 reviews: 4 for chat 117, 3 for chat 116, 2 for chats 118, 119 and 120 and a single review for chats 125 and 126. At the same time, we asked the chat participants to rank their colleagues with respect to their activity in the chat. As a third way to evaluate the results of the proposed system and assessment factors, we also considered the scores provided by PolyCAFe. The average correlation and the standard deviation between the overall results of our system and the ones provided by the tutors, the students and PolyCAFe are shown in the overall column from Table 2 (the bolded values are the best ones, while the ones in italics are the worst). The results for each individual factor are shown on their own column in order to be able to determine their influence for the overall score.

Table 2. Spearman's rank correlation between the application's results, the tutors, the students and PolyCAFe

Chats	Correlation with…	involvement		knowledge		innovation		overall	
		single	multi	single	multi	single	multi	single	multi
116	tutor-overall	0.5	0.3	0.9	0.9	-0.4	-0.4	0.9	0.9
	PolyCAFe	0.7	0.4	1	1	-0.3	-0.3	1	1
	stud-overall	0.5	0.3	0.9	0.9	-0.4	-0.4	0.9	0.9
117	tutor-overall	0.7	-0.5	0.9	0.6	-0.5	-0.5	0.8	0.8
	PolyCAFe	0.7	-0.5	0.9	0.6	-0.5	-0.5	0.8	0.8
	stud-overall	0.7	-0.5	0.9	0.6	-0.5	-0.5	0.8	0.8
118	tutor-overall	0.7	0.2	0.9	1	-0.6	-0.6	0.9	0.9
	PolyCAFe	0.5	-0.1	0.7	0.9	-0.7	-0.7	1	1
	stud-overall	0.7	0.2	0.9	1	-0.6	-0.6	0.9	0.9
119	tutor-overall	1	1	0.9	0.9	-0.7	-0.7	0.9	0.9
	PolyCAFe	1	1	0.9	0.9	-0.7	-0.7	0.9	0.9
	stud-overall	0.5	0.5	0.8	0.8	-0.9	-0.9	0.8	0.8
120	tutor-overall	0.5	0.5	1	1	-0.4	-0.4	1	1
	PolyCAFe	0.6	0.6	0.9	0.9	-0.3	-0.3	0.9	0.9
	stud-overall	0.3	0.3	0.7	0.7	-0.5	-0.5	0.7	0.7
125	tutor-overall	0.5	0.5	1	0.9	-0.6	-0.6	0.9	0.9
	PolyCAFe	0.5	0.5	1	0.9	-0.6	-0.6	0.9	0.9
	stud-overall	0.3	0.3	0.9	1	-0.8	-0.8	1	1
126	tutor-overall	0.3	0.6	0.7	0.7	-1	-1	0.9	0.9
	PolyCAFe	0.6	0.8	0.9	0.9	-0.9	-0.9	1	1
	stud-overall	0.7	0.4	1	1	-0.3	-0.3	1	1
AVG	tutor-overall	0.6	*0.37*	**0.9**	0.86	-0.6	-0.6	**0.9**	**0.9**
	PolyCAFe	0.66	*0.39*	0.9	0.87	-0.57	-0.57	**0.93**	**0.93**
	stud-overall	0.47	*0.24*	0.83	0.81	-0.67	-0.67	**0.86**	**0.86**
DEV	tutor-overall	0.22	*0.46*	0.1	0.15	0.21	0.21	**0.06**	**0.06**
	PolyCAFe	0.17	*0.52*	0.1	0.13	0.22	0.22	**0.08**	**0.08**
	stud-overall	0.18	*0.36*	**0.1**	0.16	0.23	0.23	**0.1**	**0.1**

The overall correlation was lower than expected, having an average of 0.60 with the tutors, 0.66 with PolyCAFe and 0.47 with the students for the single chat analysis and even worse for the multi-chat evaluation - 0.37 for tutors, 0.39 for PolyCAFe and 0.24 for students. Moreover, it can also be easily observed that very large values for the standard deviation have been obtained showing that the overall score is not very suitable for assessing the contribution of the participants. These results are also poor when compared to the correlation between tutors-students (average 0.871, standard deviation 0.19), tutors-PolyCAFe (average 0.94, standard deviation 0.05) and students-PolyCAFe (average 0.85, standard deviation 0.16).

Therefore, we have tried to see which one of the three components of the overall score introduced these problems. This is why we ended up computing the correlation between the gold standard provided by the tutors, students and PolyCAFe with our results for each of the three components of the overall score: involvement, knowledge and innovation.

The correlation with the involvement heuristic proved to be extremely good, with an average of 0.90 with tutors and PolyCAFe and 0.83 with students in the case of single chat analysis and 0.86, 0.87 and 0.81 in the case of multi-chat analysis. The standard deviations are also good as they have very low values: 0.10 in the case of single chat analysis and 0.15, 0.13 and 0.16 for multi-chat analysis. The innovation heuristic also seemed to be extremely well correlated with the gold standard, obtaining an average correlation of 0.90 with tutors, 0.93 with PolyCAFe and 0.86 with students and having a standard deviation of 0.06, 0.08 and 0.10. These results were the same in both cases of the single and multi-chat analysis.

Finally, the correlation between the gold standard and the knowledge heuristic provided us a great surprise: most of the time, our results proved to be anti-correlated with the gold standard. For this heuristic, both the single and multi-chat analysis had the same results: the average was -0.60 with tutors, -0.57 with PolyCAFe and -0.67 with students, while the standard deviation was 0.21, 0.22 and 0.23.

Considering this strong anti-correlation for this factor, we believe that the main problem with our overall results was provided by the way we considered knowledge in the final score. Another important information that we extracted from this evaluation of the assessment factors was that knowledge was less correlated with the golden standard even if we would have considered the absolute value of the correlation value (around 0.60 compared to 0.90 in the case of the other two heuristics). This tells us that even if we consider the inverse ranking provided by knowledge metric instead of the current ranking, the results will still be worse than if simply ignoring this heuristic.

Encouraged by the above finding, we continued with the 5 factors influencing the involvement heuristic in order to identify which are the most important ones and which can be ignored. The average and standard deviation values are presented in Table 3 (again, the best values are bolded, while the worst are emphasized with italics characters). Based on these results, we identified another heuristic that was anti-correlated with the gold standard: *Persistence*. Besides this heuristic, we also identified that *Activity* is not well correlated with the manual annotation (average correlation was 0.24, 0.37 and 0.37 while the standard deviation was 0.51, 0.45 and 0.57).

In conclusion, from the initial 5 heuristics considered together to characterize the participants' involvement, only 3 actually provide important evidence to motivate their use: Number of replies, Absence and Repetition. From these 3, the most important factor seems to be the Absence, while the least important is the Number of replies.

Table 3. Spearman's rank correlation between the gold standard results and the factors influencing Involvement

Factors	Average			Standard Deviation		
	tutor - overall	PolyCAFe	Stud - overall	tutor - overall	PolyCAFe	Stud - overall
Replies	0.828671	0.757143	0.685714	**0.213809**	**0.21492**	0.260951
Activity	*0.242857*	*0.371429*	*0.371429*	*0.512696*	*0.453557*	*0.576525*
Absence	**0.857143**	**0.8**	0.714286	0.229907	0.223607	0.291139
Persistence	-0.6	-0.51429	-0.42857	0.326599	0.34365	0.415188
Repetitions	0.771429	0.757143	**0.814286**	0.256348	0.243975	**0.226779**

Regarding the content of the 7 conversations, we did not have a comparative evaluation of the considered chats, neither from students nor from tutors. Therefore, we used the analysis provided by PolyCAFe [7] and compared the results obtained by us with the ones provided by PolyCAFe. Since we only had the comparative results for the four imposed concepts – chat, blog, forum and wiki – (see Table 4) while PolyCAFe offered an overall score, we identified for each of these concepts the ranking of the 7 chats and after that we considered the sum of these ranks for every conversation (column "Sum of ranks" in Table 4). Afterwards we re-rank the chats according to this sum, considering that the smaller value represents a higher ranking, and this way we computed the overall rank for the whole set of conversations in the corpus. What is extremely interesting is that the Spearman's Rank Correlation between our overall ranks and the ones provided by PolyCAFe was 0.964, an extremely high correlation, especially considering the fact that the content was evaluated in two different ways: in PolyCAFe based on coherence and degree of coverage using Latent Semantic Analysis (LSA) was used for content evaluation, while in our system we considered the lexical chains built on top of WordNet.

Table 4. Content evaluation and comparison to PolyCAFe

Rhythmicity of concepts / Chats	chat	Rank chat	blog	Rank blog	forum	Rank forum	wiki	Rank wiki	Sum of ranks	Overall Rank	PolyCAFe
116	4.3	6	3.41	6	3.79	5	4.96	2	13	3	2
117	5.23	3	4.95	2	4.79	3	4.75	3	21	6	6
118	6.05	2	4.49	3	4.12	4	4.72	4	19	4	4
119	5.05	4	3.93	5	3.43	6	3.33	6	11	2	3
120	4.58	5	3.94	4	4.84	2	6.1	1	20	5	5
125	4.05	7	3.29	7	3.11	7	2.66	7	4	1	1
126	6.08	1	5.07	1	6.47	1	4.59	5	24	7	7

5 Conclusions

In this paper we presented an application that can be used for analyzing the participants to a chat conversation and also to score them according to their involvement, knowledge, innovation introduced in the chat. Moreover, all these factors have been combined to determine their overall activity. The application also allows the comparative analysis of the participants from a series of conversations that debate the same subject, considering that the relevant topics to be debated are provided by the tutors.

We have shown that the overall score computed by the application is not very reliable especially when compared to other systems. However, when analyzing each component used to compute the overall score, we have found that some of the heuristics perform quite well and that the overall results are affected by only a single factor. The heuristics that proved to work best were the participants' innovation and involvement, while the one used to assess knowledge was either poorly designed or poorly interpreted. Moreover, for computing the involvement heuristic, we found out that it is better to avoid considering the participants' persistence in the chat and their activity counted as the average number of characters per uttered reply as they are also poorly correlated with the manual annotation.

However, the most important result in this paper is the methodology of how to identify which heuristics work best and which are the ones that should be avoided if a combined score should be computed by a given application for assessing CSCL chats or the results of other tasks.

Acknowledgement. This research was supported by project No.264207, ERRIC-Empowering Romanian Research on Intelligent Information Technologies/FP7-REGPOT-2010-1.

References

1. Baker, M. J. and Lund, K. Flexibly structuring the interaction in a CSCL environment. Paper presented at the European Conference on Artificial Intelligence in Education (1996)
2. Bakhtin M.M. Problems of Dostoevsky's Poetics, Ardis (1973)
3. Chiru, C., Cojocaru, V. Trausan-Matu, S., Rebedea, T. and Mihaila, D. Repetition and Rhythmicity Based Assessment for Chat Conversation. In: ISMIS 2011, LNCS 6804, Springer, pp 513-523, ISBN 978-3-642-21915-3 (2011)
4. Elsner, M. and Charniak, E. Disentangling Chat. In: Computational Linguistics, 36(3), pp. 389-409 (2010)
5. Jurafsky, D., and Martin, J. H. Speech and Language Processing: An Introduction to NLP, Computational Linguistics and Speech Recognition. 2nd edition. Prentice-Hall (2009)
6. Muhlpfordt, M. and Wessner, M. Explicit referencing in chat supports collaborative learning. Paper presented at the Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!, Taipei, Taiwan (2005)
7. Rebedea, T., Dascălu, M., Trausan-Matu, Armitt, G., and Chiru, C. Automatic Assessment of Collaborative Chat Conversations with PolyCAFe. In: Proceedings of ECTEL 2011, LNCS 6964, Springer, pp. 299-312 (2011)

8. Schegloff, E. Discourse as an Interactional Achievement: Some Uses of 'uh huh' and Other Things That Come Between Sentences. In: Analyzing Discourse: Text and Talk (Georgetown University Roundtable on Languages and Linguistics 1981), D. Tannen, Ed. Georgetown University Press, pp. 71-93 (1981)
9. Simpson, J. Conversational floors in synchronous text-based CMC discourse. In: Discourse Studies, 7(3), pp. 337-361 (2005)
10. Stahl, G. Group cognition. Computer support for building collaborative knowledge. Cambridge: MIT Press (2006)
11. Stahl, G. Studying Virtual Math Teams. New York: Springer (2009)
12. Tannen, D. Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse. Cambridge University Press (1989)
13. Trausan-Matu, S. and Rebedea, T. A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants. In: A. Gelbukh (Ed.), CICLing 2010, LNCS 6008, pp. 354–363, Springer (2010)
14. Voloshinov, V. N. Marxism and the Philosophy of Language. New York Seminar Press (1973)

Research Evidence on the Impact of Technology-Enhanced Collaboration Scripts on Learning

A Contribution toward a Script Theory of Guidance in CSCL

Stavros Demetriadis

PO Box 114, Department of Informatics, Aristotle Univ. Campus, 54124, Thessaloniki, Greece
sdemetri@csd.auth.gr

Abstract. This work summarizes key findings of current research in the CSCL domain as a contribution to the construction of a consistent theoretical framework that encapsulates the multifaceted aspects of scripted collaboration (proposed as a "script theory of guidance"). Based on field research evidence, four principles are suggested and discussed, namely "Script Appropriation", "Role as Script Configuration Cue", "Coercion on Explicitness" and "Task to foster Transactivity" principles. The key conceptualization set forth by these principles is that: (a) learners' implementation of a script is a socially negotiated process; (b) a role assigned to a student may also affect other partners' internal script induction/configuration; (c) teachers' interventions to coerce students on being explicit and implement 'transactivity fostering' tasks may have a significant impact on learning outcomes. These conclusions are also discussed from the perspective of system designer as guidelines toward developing research-informed efficient technology tools for scripted collaboration.

Keywords: Collaboration scripts, Script theory of guidance, CSCL, Technology-enhanced scripting, Transactivity.

1 Introduction

This work concisely presents and comprehensively discusses the outcomes of a series of studies that the CLASS research group conducted to explore the impact of collaboration scripts on student learning. CLASS ("Collaborative Learning Activity Support Systems") group is a research group at the Department of Informatics, Aristotle University of Thessaloniki specializing on technology-enhanced learning systems with emphasis on scripted CSCL (Computer-Supported Collaborative Learning) systems and relevant instructional techniques (http://mlab.csd.auth.gr/index.php/en/class).

The added value of this work is that the outcomes of these studies –most of them have already been published in various international journals- are discussed in the light of a recently proposed "script theory of guidance", articulated as a set of principles compiling and founded on available research evidence relevant to scripted collaboration [1]. Through this perspective, this work is considered as a contribution to the further formulation of the script theory, either by linking existing field research

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 97–110, 2013.

evidence to the theoretical construct or by highlighting specific facets of scripting practice and implementation that call for further elaboration of the theory. Additionally, the conclusions reported are discussed from the designer's perspective, boiled down to design suggestions for CSCL scripting tools (for example, script editing tools, etc.). The following sections include: (a) the theoretical background (section 2); (b) the studies compilation organized in subsections with a common theme (section 3); (c) implications and discussion (section 4).

2 Theoretical Background

2.1 Agents in a Scripted CSCL Setting

A "scripted CSCL setting" is a learning situation where learners work collaboratively (usually in small groups of 2-4 students) to accomplish a learning task, supported by some kind of digital technology. A key component of the setting is the *collaboration script* ([2, 3]), that is, a didactic scenario (usually teacher-led) imposing structure on the activity by organizing it in a sequence of phases, defining tasks, deliverables and time-schedule. The script also offers guidance to students on how to collaborate productively, mainly by assigning roles which trigger and direct peer interaction. Major agents in this setting are the following:

Learner. In a typical CSCL setting the individual learners (students) are distributed in workgroups and are guided to engage in fruitful interactions with their partners ("peer interactions"), which is considered as the main learning mechanism in a collaborative learning situation [4]. Peers are expected to: (a) develop and maintain a shared understanding of the collaborative task, (b) negotiate, realize and validate a possible solution and working path for accomplishing the task. Individual learning is expected to emerge when students activate cognitive processes triggered by peer interactions in all the above work-phases. Consequently, the level and quality of peer interactions is considered to affect extensively the level of individual learning constructed during collaboration.

Group. The group of partners (peers) is a higher level agent and the activity of a small group is the typical unit of analysis in CSCL studies, focusing both on peer interactions within the group and group interactions with other groups and agents in the setting. Thus, "group learning" is a term encompassing both the analysis of peer interactions and the learning outcomes attributed to the group and documented through the development and assessment of joint deliverables.

Script. As students do not usually engage in productive peer interactions when left unguided, researchers came up with specific techniques to script the collaborative activity [5]. The notion of scripting [6] is strongly connected to the externally (teacher-led) imposed structure and guidance to peers for implementing the collaborative task. A collaboration script organizes the activity into a sequence of phases, assigns student roles and tasks, and explicates the specificities of role playing and task development, with the ultimate goal of triggering and guiding peer interactions. Overall, a collaboration script can be seen as an instructional technique aiming to increase the probability that fruitful peer interactions will occur during the activity. The potential of the scripting technique to advance both domain-specific and domain-general learning outcomes (that is, students' "transversal" skills such as argumentation, peer reviewing, etc.) is well documented by several studies in current CSCL literature (for example [5, 7, 8]). A framework for script major constituents [2] suggests that any script can be conceptualized as entailing both "*mechanisms*" and "*components*". Script components include entities as: participants, activities, roles, resources and student groups. Script mechanisms include: task distribution (how activities, roles, and resources are distributed across participants), group formation (how participants are distributed across groups, and sequencing (how both components and groups are distributed over time). Significant aspects of the script conceptual framework include, also, the following:

1. *Macro- vs. micro- scripts* (or '*scriptlets*' [1]). A macro-script provides description of the overall collaborative activity framework (for example, describing the phases, student roles and tasks of an online argumentation activity). By contrast, a micro-script (scriptlet) is a focused set of guidelines of how to practice a specific skill (for example, an argumentation scriptlet presents to students guidelines on how to construct quality arguments). A scriptlet usually is considered as a piece of knowledge worth to be internalized by peers.
2. *External vs. internal student scripts*. While teachers employ some kind of external representations to present the external script to students, individuals also hold and activate "internal" scripts, that is procedural knowledge about how to act in situations like the ones imposed by the external script. The interaction between external and internal script is very important in a scripted CSCL setting since this interaction is expected to influence the actual form of the script implemented [9].

Technology-Enhanced Scripting. In this context, the term refers to the use of digital tools and services to support, improve and extend the potential of scripting techniques in triggering, modeling, and scaffolding peer interactions that foster high quality learning conditions for group learners. Technology-enhanced scripting may refer to actions as simple and trivial as uploading a text-based script representation on a webpage for peers to review or as complex and innovative as building and evaluating script editing tools (e.g. [10]) or system architectures for standard-based data transfer among CSCL tools (e.g. [11]). Scripting has triggered CSCL researchers' interest both from the pedagogical and technological perspective. For example, Dillenbourg and Tchounikine have extensively discussed pedagogical issues of scripting [12], while a new family of technological tools have been offered to users, that can support the editing and deployment of scripted collaborative scenarios (on the web or locally) (for example, [13], [14]).

Teacher. In a CSCL setting the teacher/instructor is considered as the "orchestrator" of the collaborative activity. Orchestration is still a rather fuzzy term in the field of technology-enhanced learning. Fischer and Dillenbourg [15] define it as "*the process of productively coordinating supportive interventions across multiple learning activi-*

ties occurring at multiple social levels". When orchestrating an activity the teacher is expected to arrange students in groups, explain the script guidelines, manage implementation schedule, trigger peer interactions by assigning student roles, provide support and feedback to students, help evaluating the task products and level of learning and –in general- engage in every aspect of script operationalization.

2.2 "Script Theory of Guidance" in a Nutshell

Fischer, Kollar, Stegmann, and Wecker [1] have recently proposed a script theory of guidance for computer-supported collaborative learning. As the authors emphasize, the theory "*addresses the question how CSCL practices are shaped by dynamically re-configured internal collaboration scripts of the participating learners, …explains how internal collaboration scripts develop through participation in CSCL practices, …it prioritizes transactive over non-transactive forms of knowledge application in order to facilitate learning and …explains how external collaboration scripts modify CSCL practices and how they influence the development of internal collaboration scripts"*(p. 2)

The seven principles of the theory strongly focus on explaining the dynamic character of learner's internal collaboration scripts and clarifying the nature of interaction between internal and external collaboration scripts. This might further provide a solid basis for conceptualizing key interactions (student-student, student-teacher and student-technology interactions) within the scripted CSCL setting, from the perspective of the dynamic process of students' internal collaboration script configuration. The seven principles proposed by the theory (as presented in [1]) are as follows:

1. *(Guidance) Internal script guidance principle*: When participating in a CSCL practice, the learner's understanding of and acting in this situation is guided by dynamically configured and re-configured internal collaboration scripts consisting of play, scene, scriptlet, and role components.
2. *(Configuration) Internal script configuration principle*: How an internal collaboration script is dynamically configured by a learner from the available components to guide the processing of a given situation, is influenced by the learner's set of goals and by perceived situational characteristics.
3. *(Induction) Internal script induction principle*: If a learner participates in an initially unfamiliar CSCL practice, then he or she builds a new configuration of already available internal script components and, through repeated application of this configuration of internal script components, develops new higher-level components (play, scene, or role) that organize the subordinate components (scenes, roles and scriptlets) for this CSCL practice.
4. *(Reconfiguration) Internal script reconfiguration principle*: If a learner's employed internal collaboration script (i.e., a configuration of internal script components) does not lead to understanding or successful actions in a CSCL practice, the internal collaboration script configuration is likely to be modified.

5. *Transactivity principle*: The more a given CSCL practice requires the transactive application of knowledge, the better this knowledge is learned through participation in this CSCL practice.
6. *External script guidance principle*: External collaboration scripts enable learners to engage in an instance of a CSCL practice at a level beyond what they would be able to without an external collaboration script either by inhibiting the automated use of internal script components or by inducing the application of internal script components that are not yet organized by a specific higher-level script component.
7. *Optimal external scripting level principle*: An external collaboration script is most effective for knowledge acquisition if it is directed at the highest possible hierarchical level of internal collaboration script components for which subordinate components are already available to the learner.

The first four principles (Guidance/Configuration/Induction/Reconfiguration) provide complementary views of the plasticity and flexibility of students' internal collaboration scripts. Essentially, these principles conclude that learners in a collaborative learning situation configure and activate various forms of internal collaboration scripts based on their goals, perception and interpretation of the current situation and also previous experiences. These constructs support learners in understanding the whole setting (play), the specific phases of the setting (scenes), the "how to" in performing a task (scriptlets) and the personal duties and course of action (role). Internally configured scripts guide learners in deciding about and implementing an optimal collaboration strategy. In case the configuration fails as the activity proceeds (for example, objectives are not achieved), the internal script form can be dynamically reconfigured and a different configuration be implemented affecting either the lower parts of the hierarchy (scriptlet or role) or the understanding of the overall activity (play, scenes). The external script is expected to inhibit or induce learners' actions and operations targeting at the right level of script hierarchy, for which the learner does not possess adequate internal script components (principles 6 and 7 in the above list).

Transactivity. The proposed theory emphasizes the issue of "transactivity" suggesting that the more productive collaboration scripts are those that promote transactive forms of discource. Transactivity ([16, 17]) is usually defined as "*reasoning operating on the reasoning of the other*", indicating to what extent "*learners build on, relate to, and refer to what their learning partners have said or written during the interaction*" [18]. From the educator's viewpoint this principle provides an important link to cognitive theories of learning, clarifying also what "productive" might refer to in the expression "productive collaboration script". Several studies have provided evidence that the more learners engage in transactive forms of dialogue the more they benefit from the collaborative activity, as the transactive nature of discourse increases the probability that learners trigger cognitive activity fostering individual knowledge construction [17, 18, 19, 20].

3 Research Studies

Working in the field of technology-enhanced scripted collaboration, our research group has published a number of studies that can contribute to the further formulation of the script theory framework. In this section, the context, the key outcomes and the implications of each study are concisely presented.

3.1 RS1: Script Appropriation

Context. In [21] we explored how group self-organization processes affected the actually implemented script during script run-time. Two groups of students studied material using a web environment designed for supporting case-based learning. Students in control group (Nc=20) studied the cases and answered a set of relevant questions individually. Students in the treatment group (Nt=20) worked in dyads (CSCL dyads), being scaffolded by a collaboration script that orchestrated the collaboration as a peer review activity. Each student in a CSCL dyad had to study initially the case-based material and provide answers individually to the set of questions. Then each partner's answers became available to their peer for review. Finally, the peers had to discuss on their reviews, identify possible different viewpoints and reach an agreement, producing and submitting their joint final answers. They were also asked to explain their decisions on how they handled their different viewpoints and produced the final deliverable. Students were allowed to communicate any way they liked during collaboration (e.g., face-to-face meeting, phone call, email etc.)

Results. Statistical analysis revealed no significant differences in the learning outcomes of the two groups (control vs. CSCL). Qualitative analysis (based on students' interviews, field observations and log file analysis) indicated that CSCL dyads self-organization resulted to a broad range of script implementation ranging from full conformance to partial violation of the script guidelines. Analysis further revealed four different versions of actually implemented collaboration scripts, depending on the degree of peer interaction during the guided peer review process ("ideal script", "moderate interaction", "weak interaction" and "no interaction").

Implications. The study results strongly suggest that scripted collaboration enacts a social process of group self-organization during which a central issue is the uptake of teacher-imposed script. Students in a group may negotiate to develop a shared understanding and adjust the teacher-imposed external script depending on their motivation for engagement, on their understanding of the learning benefits from script implementation (relevant to students' metacognitive skills) and on script coerciveness directly related to teacher's detailed script design (in general, a more detailed script design results to a more coercive form of the script). This conclusion corroborates the remark that the "ideal" script (as conceptualized by the teacher) can be different eventually, being compared to the form of the script actually implemented ('actual' script) [22].

We call this process "script appropriation". Appropriation, a key idea in the milieu of socio-cultural learning theories, refers to the process of learner's transforming a cultural construct (concept, process, tool), adapting it to one's own understanding, needs and intentions [44]. We use the term to emphasize that peers negotiate and develop their own script form, which, in turn, reflects their interpretation of the suggested activity. We also maintain that it is not a process implying suboptimal teacher script design, but a process inherent to the teaching/learning process and "*to the notion of task and human activity*" ([23], p. 36). Since learners appropriate the external script, the configuration of internal collaboration scripts to enact in a specific situation is not a passive, individual-based and straightforward process but rather an active, socially enacted and negotiated process.

3.2 RS2: Coercion on Explicitness

Context. In [24] we investigated how making the script more coercive can influence student script appropriation and –subsequently- learning outcomes. This study explored the impact of the proposed "Make It Explicit!" technique on student learning when participating in scripted collaborative activities. The method posits that when asking students to proactively articulate their own positions explicitly (that is, in written form), then conditions for improved peer interaction are fostered for the subsequent phases of collaboration. Forty two (42) students worked in dyads on a peer review scripted task supported by an appropriately designed web-based system. First, students were asked to individually study the same material and provide answers to relevant questions, then review their peer's contribution and, finally, collaborate again in dyads to provide their revised joint answers to the same questions. Dyads were randomly assigned in two conditions: Low Coercion (LC) and High Coercion (HC). In the HC condition, students were required to explicitly produce (write and submit through the system) their reviews, prior to the task of interacting to jointly develop their final answers to the questions. In the LC condition, instead, submitting the individual review in written form was optional.

Results. By collecting and analyzing both quantitative (pre- and post-test scores, measures of individual and group learning) and qualitative (interviews, log files) data, it was shown that students in the HC condition applied better collaboration patterns (in terms of peer interaction) and outperformed students in the LC condition in acquiring domain conceptual knowledge. These differences occurred mainly because of the imposed requirement on the HC group that peers make explicit their comments and opinions regarding partner's contributions (written reviews). First, submitting the reviews in writing can be beneficial for learners, as the process of producing an external representation help them articulate their own thinking and understanding [25, 26]. Second, by making their opinions explicit, students are informed about peer's opinions and understanding, and enter the discussion aware of each other's position and able to easily identify different views and disagreements. In this case, a written review fosters a more transactive form of peer dialogue, that is, supports students to efficiently reason on each other's positions and contributions. Both these points came up in the interviews we held with the HC group students, suggesting that the benefits of the written reviews were actually acknowledged. By contrast, in the LC condition 9 out of the 10 dyads chose not to submit their comments in the system. This decision altered significantly the learning conditions for the LC students, who stated that they also performed a review process, but they preferred to submit their comments directly to their peers in an unstructured way.

Implications. This study has shown that peers can be effectively guided to collaboratively elaborate on learning material and attain more knowledge with a high-coercion script than promotes the "explicitness" in peer contributions. The "Make it Explicit!" method suggests that increasing script coercion by asking students to become explicit in their positions (thinking, arguments, statements, etc.) increases the probability of improved learning outcomes ("*Coercion on explicitness*" principle). This perspective is in line with other studies reporting on the benefit of asking peers to make explicit key aspects of their shared task conceptualization [27]. We attribute the identified beneficial impact on the combined effect of the transactive form of dialogue fostered by peers' explicit representations and the fact that a writing task in itself has a positive impact on learning. We acknowledge that further research is needed to clarify this point.

3.3 RS3: Role as Script Configuration Cue

Context. A series of our studies analyzed the hypothesized beneficial impact of online scripted collaboration in the domain of algorithm learning. In [28] 24 students were randomly distributed in two groups (control: N=12, 6 dyads; treatment: N=12, 6 dyads) and were assigned a problem-solving task on LL(1) parsing algorithm. Students worked online in an integrated environment using a chat tool to communicate and an AV (algorithm visualization) software tool to visualize their proposed solutions. The control dyads were given simple instructions explaining the task questions that the peers had to collaboratively answer. The treatment dyads were additionally guided in their activity by a peer tutoring script [29], which reciprocally assigned the role of tutor and tutee to the two students, providing hints to the tutor on how to pose a task-relevant issue which would help peers to deepen their understanding of the task and advance their proposed solution. A second follow-up study of similar design was conducted with 26 participants (control: N=12, 6 dyads; treatment: N=14, 7 dyads) [30], while in a third study we applied the same design but with students working in triads this time (control: N=24, 8 triads; treatment: N=24, 8 triads) [31].

Results. The results were beneficial for scripted condition in all three studies. Discourse analysis based on an appropriately extended IBIS model [32] and lag-sequential analysis (LsA) [33] indicated that, as expected, the collaboration script enhanced the on-task peer interaction and the intrinsic feedback [34] that peers receive from interacting with the AV system. Thus, the scripted approach helped avoiding the suboptimal peer interaction patterns typically reported in the literature [35]. Furthermore, statistical analysis applied on students' post-tests, revealed that the observed script impact on peer interaction had also an impact on individual learning outcomes, significantly improving the post-test performance of treatment students [30, 31]. Overall, the script guided students to engage in a more transactive form of dialogue that eventually resulted in improved individual learning outcomes.

An interesting issue came up when analyzing the interaction in the case of triads we focused on how peers collaborate when a student in the triad (S1) poses an issue (either during the natural peer discussion in the unscripted condition or when scripted to play the role of peer tutor in the scripted condition). At this point the other two students in the triad (S2, S3) were unguided on how to collaborate (either in the scripted or unscripted condition). Our data indicated that students S2 and S3 exhibited a different behavior depending on the condition. Treatment students demonstrated a more active S2-S3 interaction before answering the issue raised by S1. That is, their actions were more oriented to first discuss the issue between them and then provide a possible answer to their S1 partner. By contrast, control students reaction (either S2 or S3) was toward directly addressing S1 on the raised issue, without necessarily including the third peer in this discussion.

Implication. As the only difference between the two conditions in the analyzed situation, was the tutor role assignment to student S1, the conclusion, from this latter remark, is that a role assigned to a student may operate as a cue for partners to activate appropriate roles and inner collaboration scripts (provided, of course, that they already possess such scripts). As Kollar, Fischer and Slotta [36] explain, students situated in a collaboration process bring in also their internal scripts as their personal process related knowledge, which defines students' behavior and interpretation of the process. Activation of inner scripts may, in turn, have an impact on peer discourse, as described above. In the treatment condition whenever student S1 raised an issue playing the 'tutor' role, the other two students (S2 and S3) identified themselves as 'tutees' establishing a peer mentality and activating an internal script that guided them to first discuss the issue and then provide an answer to S1 (this was evident as more intense S2-S3 interaction in discourse analysis). By contrast, students in the control group had no reason to initiate an intense S2-S3 interaction since S1 was not socially differentiated. Directly addressing S1 on the raised issue was simply considered as addressing a peer during collaboration. The control students equally relied on their procedural knowledge (internal script) about collaboration, but the interaction within the triad was different because of not any role assigned. We consider this evidence as supporting the "*Role as script configuration cue*" principle during scripted activity.

3.4 RS4: On the Role of Fading vs. Peer Monitoring Techniques

Context. In another series of our studies we focused on the impact of two important scripting techniques in argumentation, namely 'fading' and 'peer monitoring'. The fading technique refers to the gradual withdrawal (fade-out) of scaffolds, initially supporting students in skill development, in order for students to practice the acquired skills on their own capacity. The impact of the fading technique has been explored in closeness with the concept of 'scriptlet', which denotes the procedural knowledge ('how to' knowledge) of sequencing the steps in a –usually- small scale activity and achieving a particular objective. For example, knowing the sequence of constructing an argument (i.e state a claim, then provide evidence and add a warrant, that is, an inference linking the evidence to the claim) is considered as a 'scriptlet' in the framework of script theory [1].

Research has investigated the impact of fading scaffolds to leave room for students for practicing the skill unaided (e.g. [37, 38]). However, the outcomes have been conflicting and some researchers suggest that the impact of fading is enhanced when the peer-monitoring technique is additionally implemented during the activity (e.g. [37]). Peer monitoring (PM) refers to a situation in which one learning partner can monitor the contributions of the other and provide focused feedback on his performance. It is expected that implementing "peer-monitoring" will guide students to recall both domain-general and domain-specific details and provide useful corrective feedback to their peer [39]. The reader should note at this point that the peer monitoring technique, when seen from the perspective of transactivity, is a technique that increases the transactive form of dialogue, since it triggers and encourages reasoning on peer articulations. To evaluate the impact of fading and peer monitoring we developed a web tool ('iArgue') for implementing online argumentation activities with these two techniques [40]. In a first study, thirty four (34) students collaborated remotely in dyads (in lab conditions) on a task guided by a scriptlet for argument construction [41]. The dyads were distributed in (a) the control group (continuous scriptlet support) (N=10); (b) the fading group (the scriptlet faded out after some student posts) (N=10), and (c) the peer-monitoring group (continuous scriptlet support and, additionally, prompting and guiding the students to monitor their peer contributions; thus, peer monitoring was compulsory) (N=14). In a second follow-up study of similar design, we explored whether we could replicate the outcomes of the first study simply by suggesting and encouraging students to implement peer monitoring [42]. Fifty (50) students participated, collaborating also in dyads which were randomly distributed in the same three conditions: (a) control (N=14); (b) fading (N=14), and (c) peer-monitoring (N=22) (PM was suggested but was not compulsory).

Results. Results of the first study (PM compulsory) revealed that students in the peer-monitoring condition outperformed those in both the control group and the fading group in domain-specific knowledge acquisition post-test items. Available evidence support the view that enriching the argumentation activity with the peer-monitoring technique can substantially improve learning outcomes, while simply fading-out the scriptlet component does not seem to improve student learning in any aspect. Results of the second study [42] indicate that -although there is a tendency in favor of the PM group- differences in learning outcomes do not reach statistical significance this time. Analysis of online discussions log files made evident that students did not consistently applied the peer monitoring technique as in the first study.

Implications. Concluding from the two studies we argue that peer monitoring during argumentation has a positive impact on student domain-specific learning. Considering that PM enhances the transactive form of dialogue, we interpret this outcome as additional evidence in favor of engaging students in 'transactivity fostering' tasks ("*Task to foster Transactivity*" principle). However, it is also clear that when the transactivity fostering task is not compulsory then the expected benefits may be significantly weakened, as students with low motivation might not engage in the task and do not practice the suggested interaction. One can go further generalizing, stating that instructors can improve collaborative learning outcomes by increasing script coercion in ways that a transactive form of the discource is supported. We argue that this is a point to be seriously considered when practicing scripting.

4 Implications for Educators and Technology Designers

The principles presented in the preceding sections can be summarized as follows:

- "*Script Appropriation*": External scripts initiate a student group appropriation process, as a social process entailing shared understanding development, negotiation and group self-organization.
- "*Coercion on Explicitness*": External scripts that compel students to provide written external representations of their thinking should be expected to result in better learning outcomes, when other principles are not violated (for example the 'optimal external scripting level' principle).
- "*Role as Script Configuration Cue*": Assigning a role to a partner may trigger other teammates as well to activate relevant inner script components, which, in turn, could have substantial impact on peer interaction.
- "*Task to foster Transactivity*": Engaging students in 'transactivity fostering' tasks (such as peer monitoring) should be expected to result in better learning outcomes, when other principles are not violated (for example the 'optimal external scripting level' principle).

The "script appropriation" principle provides a perspective in conceptualizing the external-internal script interaction, emphasizing the social process through which groups adopt external guidance, develop their "own" script view and configure/activate/re-configure respective forms of internal scripts. The implication for educators is that the appropriation process should be anticipated as an essential part of the scripted activity, it should be encouraged, so that student groups explicitly state and negotiate their (mis-) understandings, and also be skillfully managed when key aspects of the script should be implemented intact or else the didactical power of the script may be severely harmed [12]. The implication for designers, in defining specifications of technology tools for scripted collaboration, is that the appropriation process should be considered and supported during script design and runtime phases. A possible idea is that script editing tools provide the functionality for designing online discussion activities where the various aspects of the proposed script (phases, roles etc.) could be discussed and negotiated in a structured manner. Also, editing tools could provide notation for emphasizing core non-negotiable components of the script (possibly enabling also teachers to provide metacognitive level support to clarify the reasons of non-negotiability).

The "role as script configuration cue" principle provides incentive toward developing role-relevant functionalities (such as a role-based inventory and role editing operations) for flexible role setting and assignment while editing the script. Efficient technology tools for role management in script design (for example, creating a role prototype, link the role play to a specific scriptlet, select which role to assign to a peer, etc.) can positively affect the configuration and enactment of appropriate student internal scripts during script runtime and further support teachers in their own role as 'directors'/'orchestrators' of the activity.

Finally, the "coercion on explicitness" and "task to foster transactivity" principles could possibly be merged to form the "coercion on student actions and tasks that foster transactivity" principle. This means that script features (activities, roles, tasks, etc.) that promote the transactive form of dialogue should be strongly encouraged and also be required from students when the situation permits. Currently, we prefer not to merge or restate the two proposed principles as the learning benefits associated with "explicitness" cannot be univocally attributed only to transactivity and further research is needed on this aspect. From the perspective of the designer these two principles suggest that editing tools for scripted collaboration should offer to users the functionality of integrating and highlighting transactivity-fostering script features (tasks etc.) in the activity workflow, considering also the availability of regulative mechanisms that enable triggering the task either in compulsory or elective mode.

Overall, these principles are presented as research-based inferences that advance our current understanding of the dynamic nature of collaboration scripting, enhance the script theory framework and provide also the basis for enriching the design space and editing capabilities of technology tools to support scripted collaborative problem-solving activities. The specifications of script editing tools, that, have been typically integrating commonly implemented design patterns (e.g. [43]), can be updated to include also appropriation support, role editing capabilities and transactivity support functionality as proposed by the presented principles.

References

1. Fischer, F., Kollar, I., Stegmann, K., Wecker, C.: Toward a script theory of guidance in computer-supported collaborative learning. Educational Psychologist 48(1), 56–66 (2013)
2. Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., Fischer, F.: Specifying computer-supported collaboration scripts. International Journal of Computer-Supported Collaborative Learning 2(2-3), 211–224 (2007)
3. Weinberger, A., Ertl, B., Fischer, F., Mandl, H.: Epistemic and social scripts in computer-supported collaborative learning. Instructional Science 33(1), 1–30 (2005)
4. Dillenbourg, P.: What do you mean by collaborative learning? In: Dillenbourg, P. (ed.) Col-laborative-learning: Cognitive and Computational Approaches, pp. 1–19. Elsevier, Oxford (1999)
5. Weinberger, A., Stegmann, K., Fischer, F.: Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). Computers in Human Behavior 26(4), 506–515 (2010)
6. O'Donnell, A.M., Dansereau, D.F.: Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In: Hertz-Lazarowitz, R., Miller, N. (eds.) Interaction in Cooperative Groups. The Theoretical Anatomy of Group Learning, pp. 120–141. Cambridge University Press, Cambridge (1992)
7. Stegmann, K., Weinberger, A., Fischer, F.: Facilitating argumentative knowledge construction with computer-supported collaboration scripts. International Journal of Computer-Supported Collaborative Learning 2(4), 421–447 (2007)
8. Weinberger, A., Stegmann, K., Fischer, F., Mandl, H.: Scripting argumentative knowledge construction in computer-supported learning environments. In: Fischer, F., Kollar, I., Mandl, H., Haake, J.M. (eds.) Scripting Computer-supported Collaborative Learning, pp. 191–211. Springer, New York (2007)
9. Kollar, I., Fischer, F., Slotta, J.D.: Internal and external scripts in computer-supported collaborative inquiry learning. Learning and Instruction 17(6), 708–721 (2007)

10. Hernández-Leo, D., Villasclaras-Fernández, E.D., Jorrín-Abellán, I.M., Asensio-Pérez, J.I., Dimitriadis, Y., Ruiz-Requies, I., Rubia-Avi, B.: Collage, a Collaborative Learning Design Editor Based on Patterns Special Issue on Learning Design. JETS 9(1), 58–71 (2006)
11. Magnisalis, I., Demetriadis, S.: Modeling Adaptation Patterns with IMS-LD Specification: A Case Study as a Proof of Concept Implementation. In: Proceedings of International Conference on Intelligent Networking and Collaborative Systems (INCOS 2009), pp. 295–300 (2009)
12. Dillenbourg, P., Tchounikine, P.: Flexibility in macro-scripts for computer-supported collaborative learning. Journal of Computer Assisted Learning 23, 1–13 (2007)
13. LAMS: Learning Activity Management System 2.0., `http://wiki.lamsfoundation.org/dis-play/lamsdocs/Home`
14. Karakostas, A., Prieto-Santos, L.P., Dimitriadis, Y.: Opportunities and Challenges for Adaptive Collaborative Support in Distributed Learning Environments: Evaluating the GLUE! Suite of Tools. In: Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies, ICALT 2012 (2012)
15. Fischer, F., Dillenbourg, P.: Challenges of orchestrating computer-supported collaborative learning. Paper Presented at the 87Th Annual Meeting of the American Educational Research Association, AERA (2006)
16. Berkowitz, M.W., Gibbs, J.C.: Measuring the developmental features of moral discussion. Merrill-Palmer Quarterly 29(4), 399–410 (1983)
17. Teasley, S.D.: Talking about reasoning: How important is the peer in peer collaboration? In: Resnick, L.B., Säljö, R., Pontecorvo, C., Burge, B. (eds.) Discourse, Tools and Reasoning: Essays on Situated Cognition, pp. 361–384. Springer, Berlin (1997)
18. Noroozi, O., Weinberger, A., Biemans, H.J.A., Mulder, M., Chizari, M.: Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. Computers and Education 61(2), 59–76 (2013)
19. Chi, M.T.: Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science 1(1), 73–105 (2009)
20. Stegmann, K., Weinberger, A., Fischer, F.: Aktives Lernen durch Argu-mentieren: Evidenz für das Modell der Argumentativen Wissenskonstruktion in Online-Diskussionen (Active learning by argumentation: Evidence for the model of argumentative knowledge construction in online discussions). Unterrichtswissenschaft 39(3), 231–244 (2011)
21. Papadopoulos, P.M., Demetriadis, S.N., Stamelos, I.G.: Analyzing the Role of Students' Self-Organization in Scripted Collaboration: A Case Study. In: Proceedings of the 8th International Conference on Computer Supported Collaborative Learning – CSCL 2009, pp. 487–496 (2009)
22. Dillenbourg, P. (ed.): Framework for Integrated Learning. Deliverable D23-05-01-F of Kaleidoscope Network of Excellence (2004), `http://hal.archives-ouvertes.fr/docs/00/19/01/07/PDF/Dillenbourg-Kaleidoscope-2004.pdf` (retrieved September 11, 2008)
23. Tchounikine, P.: Computer science and educational software design – A resource for multidisciplinary work in technology enhanced learning. Springer, Berlin (2011)
24. Papadopoulos, P. M., Demetriadis, S. N., Weinberger, A.: Make It Explicit!: Improving Collaboration through Increase of Script Coercion. Journal of Computer Assisted Learning. 29(4), 383–398 (2013)
25. Jang, S.J.: A study of students' construction of science knowledge: Talk and writing in a collaborative group. Educational Research 49(1), 65–81 (2007)
26. Nückles, M., Hubner, S., Renkl, A.: Enhancing self-regulated learning by writing learning protocols. Learning and Instruction 19(3), 259–271 (2009)

27. Moguel, P., Tchounikine, P., Tricot, A.: Interfaces Leading Groups of Learners to Make Their Shared Problem-Solving Organization Explicit. IEEE Transactions on Learning Technologies 5(3), 199–212 (2012)
28. Foutsitzis, C., Demetriadis, S.: Scripted collaboration to leverage the impact of algorithm visualization tools in online learning: Results from two small scale studies. International Journal of e-Collaboration 9(1), 42–56 (2013)
29. Palinscar, A., Brown, A.: Reciprocal teaching of comprehension-fostering and comprehension monitoring activities. Cognition and Instruction 1(2), 117–175 (1984)
30. Foutsitzis, C., Demetriadis, S.: Collaboration scripts as a method for improving online algorithm learning with the use of visualization systems. Journal of Computer Assisted Learning (under revisions, 2013)
31. Foutsitzis, C., Demetriadis, S.: The role of collaboration scripts to foster transactive forms of peer interaction in online learning (in preparation)
32. Kunz, W., Rittel, H.: Issues as elements of information systems. In: Proceedings of the 5th International Conference on Web Information Systems and Technologies, pp. 198–211 (1970)
33. Erkens, G., Kanselaae, G., Prangsma, M., Jaspers, J.: Computer support for collaborative and argumentative writing. In: De Corte, E., Verschaffel, L., Entwistle, N., van Maerrienboer, J. (eds.) Powerful Learning Environments: Unraveling Basic Components and Dimensions, pp. 157–176. Elsevier, Amsterdam (2003)
34. Laurillard, D.: Pedagogical forms for mobile learning: Framing research questions. In: Pachler, N. (ed.) Mobile Learning: Towards a Research Agenda. WLE Center, IoE, London, UK (2007)
35. Liu, C., Tsai, C.: An analysis of peer interaction patterns as discoursed by on-line small group problem-solving activity. Computers & Education 50(3), 627–639 (2008)
36. Kollar, I., Fischer, F., Slotta, J.D.: Internal and external scripts in computer-supported collaborative inquiry learning. Learning and Instruction 17(6), 708–721 (2007)
37. Wecker, C., Fischer, F.: Fading instructional scripts: Preventing relapses into novice strategies by distributed monitoring. In: Gomez, K., Lyons, L., Radinsky, J. (eds.) Proceedings of the 9th International Conference of the Learning Sciences: Learning in the Disciplines, pp. 794–801. ISLS, Chicago (2010)
38. Renkl, A., Atkinson, K.R., Große, C.S.: How fading worked solution steps works – a cognitive load perspective. Instructional Science 32(1–2), 59–82 (2004)
39. King, A.: Discourse patterns for mediating peer learning. In: O'Donnell, A.M., King, A. (eds.) Cognitive Perspectives on Peer Learning, pp. 87–115. Erlbaum, Mahwah (1999)
40. Bouyias, Y., Demetriadis, S.: iArgue: a web-based argumentation environment integrating collaboration scripting support and flexible fading techniques. In: Pinkwart, N., McLaren, B.M. (eds.) Educational Technologies for Teaching Argumentation Skills, pp. 198–224. Bentham Science Publishers (2012)
41. Bouyias, Y.N., Demetriadis, S.N.: Peer-monitoring vs. micro-script fading for enhancing knowledge acquisition when learning in computer-supported argumentation environments. Computers & Education 59(2), 236–249 (2012)
42. Bouyias, Y.N., Demetriadis, S.N.: The impact of "peer monitoring" method in online argumentation as a task that fosters transactive forms of dialogue (submitted)
43. Hernández-Leo, D., Asensio-Pérez, J.I., Dimitriadis, Y.: Computational Representation of Collaborative Learning Flow Patterns Using IMS Learning Design Educational Technology & Society 8(4), 75–89 (2005)
44. Wertsch, J.V.: Mind as action. Oxford University Press (1998)

An Environment to Support Collaborative Learning by Modding

Sébastien George[1,2], Élise Lavoué[1,3], and Baptiste Monterrat[1,2]

[1] Université de Lyon, CNRS
[2] INSA-Lyon, LIRIS, UMR5205, F-69621, France
{sebastien.george,baptiste.monterrat}@insa-lyon.fr
[3] Université Jean Moulin Lyon 3, Magellan, Liris, UMR5205
elise.lavoue@univ-lyon3.fr

Abstract. In this paper, we propose an environment to support collaborative modding, as a new way to learn a subject. Modding can be defined as the activity to modify an existing game with dedicated tools. In a constructivist approach, we base our work on the assumption that modding a learning game can help learners to acquire the concepts of the subject concerned. We also think that modding in collaborative settings can help learners both to learn the subject and to learn to collaborate. We first propose a framework to support collaborative modding activities based on four components: the game, the Game Development Kit (GDK), contextual discussions and a knowledge map. We then propose an architecture that integrates these components on a unique platform. We finally present the results of a first exploratory study that demonstrates the feasibility and the interest of this approach for learning and the need for integrating collaborative tools in a unique environment.

Keywords: modding, game development kits, learning game 2.0, collaborative learning.

1 Introduction

During the last decade, new tools have emerged in the field of video games in order to allow novice developers to create their own games. These tools are called "game editors", "game factories" or "Game Development Kit" (GDK). Meanwhile, playing learning games has proven educational benefits by immerging learners into a world where they have to invest themselves intellectually and mentally to progress, to face challenges or to accomplish quests. In this article, we are interested in a new way to learn the knowledge of a domain: modding learning games. Modding can be defined as the activity to modify an existing game with dedicated tools.

As we show in the state of the art, few studies have focused on this method of learning in education. Those who were interested in this method mainly used it to learn programming. The work presented in this paper explores new uses of learning by modding and proposes a generic framework to support collaborative learning activities based on modding. The relative works on learning by modding is presented in Section 2.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 111–124, 2013.

We explain in Section 3 the learning approaches supported by the activity of creating and modifying the content of a learning game in a collaborative way. We then present the framework and the tools we propose to support this activity (Section 4). As this research is exploratory, we finally present the results of a first study with 16 participants (Section 5 and 6). This study aims to test the interest of the approach and the proposed environment. We finally conclude with discussions and further work.

2 Learning by Modding: State of the Art

2.1 Users Becoming Programmers

At the beginning of gaming, the only way to modify a video game was to access its source code. Fortunately, the emergence of new types of editing tools provided users with easy access to the core of the game. The most basic way to modify a game for a user is to access its settings. Most games allow modifying for example the display mode or the level of difficulty. During the 80s, new types of tools arrived: level editors. They have contributed to the success of Lode Runner, one of the first games that allowed users to describe by themselves the initial state of a level (Djaouti, 2011). Nowadays, players have access to many different game engines allowing them to completely change the behavior of a game to convert it into a new one.

Tavares and Roque (2007) have highlighted the advantages for a game to be designed by a lot of players mixing a lot of ideas, instead of a few professionals. According to Volk (2008), "the roles of game designer and game player are obviously not a binary one, since every level of participation can be found in the modding movement". In the same thought, we understand that video games need both professional design and fans ideas. As both roles exist, modding tools have to be adapted to these different profiles.

Web 2.0 has grown thanks to the WYSIWYG ("What You See Is What You Get") editors. By analogy, the game 2.0 is becoming accessible to all thanks to a new generation of tools, providing users with a good representation of the game. Djaouti et al. (2011) define game 2.0 as any application allowing a user to create, share and play to a game content. Within this approach, some tools have been created to simplify programming (Moshirnia, 2007). These tools aim to increase the expressiveness of the mode of representation while reducing its complexity. They are often composed of two interfaces: one to edit the initial state of a scene or of a game level, and the other to modify the behavior of the game.

The drag and drop of objects from a library to the game is an example of possibility opening up level editors to a wider audience. Interfaces that offer the possibility to modify the behavior of a game are also becoming increasingly visual, as these examples show:

- Stencyl and Flip allow users to program a game without writing any code. Some blocks ("if", "then", "Boolean" ...) and existing functions can easily be dragged, dropped and ordered in a structure corresponding to the behavior of the game.

– Warcraft III editor offers further possibilities with a representation of the rules on the shape of triggers (composed of an event, conditions and actions). So programming concepts such as loops and functions are implicit and do not require to be known by the developer.

– Kodu Game Lab and Game Develop use an even simpler way to represent the behavior of a game: rules consisting only of conditions and actions. The game engine works as if the conditions were constantly tested, and corresponding actions triggered if necessary. This system is so simple that Kodu allows children to design their own games. Kodu also allows access to the behavior editor within the game itself: to select an object leads directly to the list of the rules that define the behavior of this object.

These editors are just a few examples of those developed recently. For more details, Djaouti, Alvarez and Jessel (2010) present a study of fifteen "gaming 2.0" tools and explain how they support the design of serious games. However, Postigo (2008) has shown the limits of modding proprietary games. In such games, modders do not have the right to share their creations. That is why we mainly studied free games and free GDK for educational purposes.

2.2 Experiences of Modding in Educational Contexts

Loh and Byun (2009) have created a serious game by modding the NeverWinter Nights 2 game. They created Saving Adryanee, a game whose "in-game" objective is to create a healing potion, and whose educational goal is to teach the concepts of chemistry. This complex work was conducted by a team of four persons in just two months, thanks to the experience of those who have conducted modding projects before them. As in this study, experienced developers and beginners usually communicate using Web collaboration tools (chat, forums), or sometimes tools integrated into the game itself (conversations between avatars in online games). Strong support exists within these communities, allowing new developers to quickly overcome the problems they meet. Based on this experience, the authors show that "instead of waiting for new serious games to be made available, the students and teachers could learn to use GDKs as learning tools. For instance, all high school freshmen may be required to learn the GDK, and to use it as one of the project presentation tools throughout high school."

In several studies, modding has been used as an activity for learning programming. Students were able to create mods themselves. For instance, McAtamney, O'Shea and Mtenzi (2005) make students use the Crytek engine to model the future of their university campus. Through this experience, students have learned how to use C++, direct X and the scripting language LUA, while developing their skills in mathematics, physics, 3D design and event-driven programming. In another study, El-Nasr and Smith (2006) showed that there are different modding tools adapted to different types of learners. They experienced the Warcraft III Editor with high school students so that they learn the basics of algorithms. Students were able to create a game in just three days. They then used Web Driver and Unreal Engine 2.5 with students in Computer Science. In addition to the technical skills acquired, these experiences have enabled students to become familiar with the software development process. Indeed, the steps of a mod development (specifications, design, implementation, and test) and the steps of software development are roughly the same (Cignoni, 2001).

3 Learning Activities Based on Modding

3.1 Learning Content with a Constructivist Approach

Up to now, studies have shown that modding has mainly been used to teach computer science. We believe that modding can be used to learn in any domain. With modding, the content of a learning game could be the learning goal (Monterrat, Lavoué & George, 2012). In fact, if learning games help to acquire the knowledge contained in the game, we can assume that changing these games can be appropriate to deeply learn a subject. Oblinger (2006) explains that the educational potential of a learning game depends on the level of involvement of the player. The goal of our study is to create good conditions of involvement by allowing the player to change the game. A moddable learning game provides the user with a new way of learning with a game. This idea is clearly consistent with a situated approach of learning (Lave and Wenger, 1991), offering the user a way to build knowledge while designing part of a game.

Somehow, we want to allow a learner to play a teacher role by designing an educational game that others will use to learn. Teaching knowledge involves a deeper level of understanding, so it is a good way to learn. With modding, we can propose to the learner to play the role of the teacher to reach another level of expertise. As suggested by Loh and Byun (2009), students can be asked to create serious game mods for use by others in order to demonstrate their understanding in the selected subject. Furthermore, learners are led to seek further information by themselves, so they can acquire concepts that go beyond the knowledge of the teacher. According to the Magic Bullet model (Becker, 2011), the learners will increase the “external learning” (what learners will look outside of the game), and include it in the “mandatory learning” (knowledge which is necessary to have in order to progress in the game). Then, the evolution of the game will benefit the future players during a new learning session. In addition, this method of learning by modding allows the knowledge to evolve in order to be always up to date, like a wiki users keep faithful to the current world that is constantly evolving. Such a game can also be well adapted to different kinds of learners by being modified by the learners themselves.

3.2 Collaborating to Learn and Learning to Collaborate

According to Scacchi (2011), modding is a way to learn to work with others as a person who creates a mod trains to work in a team and to manage a group project. During the experiments of El-Nasr and Smith (2006), the students first learnt to divide the tasks among groups of two and share their skills. They then went beyond, exchanging with other groups because they understood that communicating about their project and discovering others would be beneficial. In bigger projects, modding also teaches how to manage a team and sometimes how to resolve conflicts. For example, Loh and Byun

(2009) have shown that, when they develop a mod, they cannot freely change the planned format of the game without offending the writer of the team. Moreover, when the number of people involved in a modding project grows, more errors and problems could appear but the quality of modding could be higher and a self-regulation mechanism may occur.

Ang, Zaphiris and Wilson (2005) have studied this "self-regulation mechanism" in communities that write wikis. We totally meet up with the idea that "by engaging learners to construct something meaningful and sharable with their peers, learners can learn by putting the knowledge into practice". The same phenomenon can occur during learning games modding, provided the game is designed as a collaborative media and modding is seen as a collaborative activity. In this way, we propose in this paper an approach to let the learners collaboratively modify the knowledge embedded in the game, as a wiki reader could participate to the document writing. In order to support this approach, we propose a particular framework in the next section.

4 An Environment to Support Learning by Modding

4.1 A Framework for Integrating Collaborative and Pedagogical Aspects

In this section, we propose a framework supporting collaborative modding in the context of learning activities (see Figure 1). This framework aims to help learners to acquire the knowledge of a specific domain predefined by a teacher. This framework consists of three parts: (1) the GDK associated with the game, (2) contextual discussions and (3) a knowledge map. The identification of these three main parts of the framework and their relationships are our main contribution for a computer supported collaborative learning by modding environment.

Game and GDK

Games are usually divided into several distinct areas called levels, worlds or cards, according to the game. We generically call them "scene". Other elements compose a video game: textures, sounds, characters, items, etc. We call them all by the generic term "object". Each scene of a game consists of a set of objects and a set of rules that describe their behavior. The GDK (Game Development Kit) is the tool associated with a game that allows modifying it. On the one hand, the GDK should to be simple enough for a quick start, with no programming skills required. Even if modding requires a learning step, this phase should be as short as possible so the user can work on the content of the game from the beginning. On the other hand, the GDK should be rich and powerful enough to allow the modder to modify the structure of the game in depth and to define new behaviors. A system of rules based on conditions and actions (as proposed in Game Develop) seems to be a good compromise. These rules allow manipulating the objects of the game, each object type having conditions and actions associated with it. A rule can be composed of any number of conditions. If there are none, the action is triggered continuously. If a rule contains multiple conditions, they all must be checked so that the action is triggered, as if they were separated by the logical operator AND. The easiest way to perform the logical operator OR is then to put the conditions in different rules. Here are some examples:

- Condition for a sound: “Is the sound <victory sound> played now?”
- Condition for the keyboard: “Is the left arrow pressed at this time?”
- Action for a sound: “Play the sound <explosion>”.
- Action for an image: “Move the image <Avatar> 10 pixels to the left”.

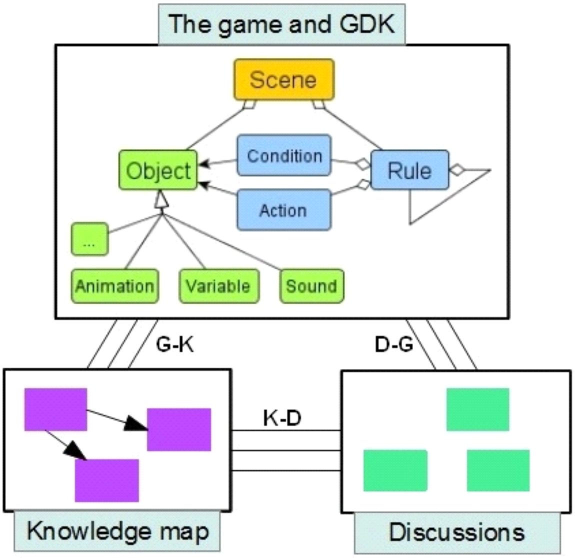


Fig. 1. Framework for a collaborative learning activity based on modding

Contextual Discussions

To work together, modders should initially be able to share ideas and opinions; to make the game emerges from collective choices. They also need to share the elements of the game so that it is actually the result of a collective work. A social platform supporting discussions and sharing of elements seems necessary. Contextual discussions have proved to be useful in educational distance situations and could be implemented for instance with contextual forums (George, 2004). That is why we propose to link contextual discussions to each elements of the game (for instance scenes or object, D-G links on Figure 1). They help students to have structured discussions on the game elements on which they are working. In the context of a modding activity, these discussions can take place in virtual spaces in order to let the learners exchange on knowledge. These contextual discussions can also be a way for the teacher to monitor, help and communicate with the students during the activity at a distance.

Knowledge Map

The knowledge map developed by the teacher contains the knowledge of the domain to learn. On the one hand, each knowledge element is linked to one or several elements of the game that allow learning it (G-K links on Figure 1). On the other hand, each knowledge element is also linked to a discussion (K-D links on Figure 1), allowing students that work on the same elements to discuss on it. Here are two examples of use in educational contexts:

– The knowledge “knowing how to apply the law of gravity” could be linked to a set of rules simulating this law in the game, and to a discussion between students about gravity relative to the size of the planet.
– The knowledge “knowing how to recognize a platypus” could be linked to the scene of a zoo and to a discussion of the differences between reptiles and mammals.

We suggest that the teacher is the supervisor of the learning activity by modding. The teacher is the expert of the domain knowledge and has to take into account the different student’s profiles when preparing the learning by modding situation. With such a framework, the teachers’ tasks are:

– To develop a learning scenario in accordance with the educational objectives.
– To create the first elements of the game that the players will modify.
– To manage and to monitor the modding activities to help learners (technically or cognitively), and ensure they meet the objectives.
– To evaluate the learners.

4.2 Integrating Tools to Support Collaborative Modding

We argue that the game and its corresponding GDK have to be integrated in a unique platform. Otherwise, switching from the game to the GDK, and from the GDK to the social platform could be a barrier for learning. That is why these 3 facets of game modding (game, GDK and discussions) have to be parts of a unique tool, as illustrated on Figure 2. Some studies have highlighted this need, for instance, McAtamney et al. (2005) noted that, in a modding activity, the students choose to use the Crytek engine particularly because it offers all the necessary tools (object database, scripting language and 3D tools) in a single environment. Furthermore that aspect matches to the definition of Game 2.0 given by Djaouti et al. (2011).

For the activity to be collaborative, it is essential that the game elements are stored on a server. Each user (client) must also have a local version of the game in order that individual tests do not affect the collaborative project. We propose to save scenes (S), objects (O) and rules (R) in a xml file that contains links to the files corresponding to the manipulated objects (images, sounds, ...). In this way, it is then easy to manage this set of files with a version control system we have included in the platform (in our case, SVN). The two main SVN commands are “update” (update the local version of the game) and “commit” (apply our changes to the collective project). Because game development involves the possibility of errors, the system should also save on the server earlier versions of the game and provide access to all these versions. Finally, a version control system can also merge two projects when necessary, for example when two learners have made changes at the same time on the same game.

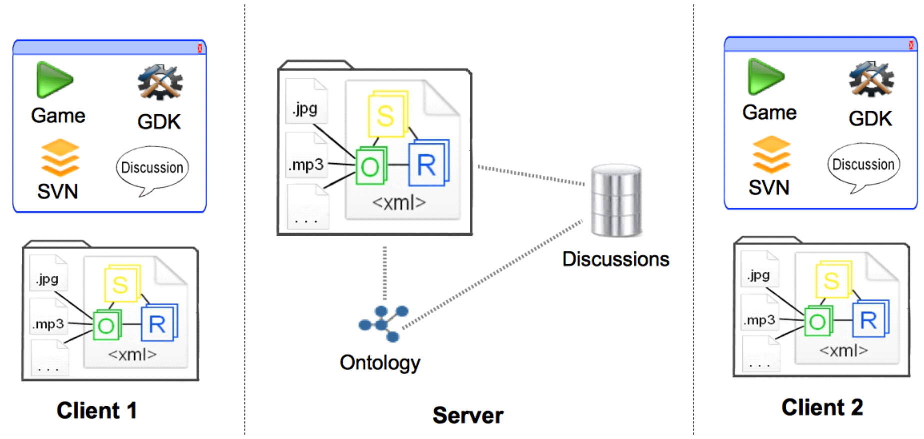


Fig. 2. Architecture of a collaborative learning by modding environment

5 Exploratory Study

5.1 Context and Participants

16 persons participated in this study. The participants were mainly students in higher education (n=14) and a few retired persons (n=2). 7 participants had prior knowledge on programming.

We proposed to the participants a prototype of modding environment. This prototype is an implementation of the system described in the previous section. According to an iterative approach, we focused this first experiment on the collaboration functionalities and we decided to not implement the knowledge map in this version. This prototype combines existing software, each addressing some of the requirements.

We chose a GDK called “Game Develop” to be the basis for our application. A wide range of conditions and actions included makes this editor expressive enough to create any kind of 2D games. In addition, the rules displayed in ordinary language makes it understandable by people without programming skills. The game was based on a central scene in which the player can find doors that give access to three other scenes. These three scenes were already created in the game, so that the students could modify them. The interface was pretty basic, in order to facilitate the use of the game by novices.

To support the discussions, we created a contextual forum available in a browser, using HTML, PHP and MySQL. In this way, each scene of the game has its own discussion thread. In parallel, the project teamwork was managed with SVN software, which allows multiple users to work remotely on common files while maintaining their versions up to date.

In order to unify the tools, a unique program, created in C language, started Game Develop (on the team project) and a browser (on the forum of the team) in one click. The participants had therefore to follow these steps at each use:

1. UPDATE (update of the local project with the contributions of other members).
2. Searching, modding, participating in discussions, etc.
3. COMMIT (update of the collective project on the server-side with the local version of the project).

5.2 Research Questions and Methodology

We used an exploratory approach in this study, as we still know little about the use of a modding environment for learning. In line with the objectives of the environment we propose, this study aims to study:

– The interest of a collaborative modding activity for learning.
– The effect of the collaborative tools we integrated in the learning activity.

The participants were asked to modify the game by using Game Develop so as to learn the Esperanto language. We divided the 16 participants into two groups: a test group and a control group. The test group was provided with the collaboration functionalities of the prototype (the contextual forum and the SVN). The members of the control group had to find by their own the means to communicate and exchange their parts of the project.

Each group was divided into two teams of 4 members (teams A and B in the control group, teams C and D in the test group). Each team has conducted the modding activity according to three main phases:

1. In presence course, by class:

- The students assisted a course on the basis of Esperanto during one hour.
- They then had a few minutes to play the game to be familiar with it.
- They finally were trained to use the GDK during an hour, with the help of the teacher.

2. During a week, the students were asked to collaborate at a distance (by group of 4) to modify the game according to the instructions. We advised them to “mod” during about 1:30 hour, with the possibility to do it in several times.
3. After this modding phase, the students of each team could play the game modded by another team in the same group (test or control).

We collected several data on this activity:

- The game that has been developed by each team. The teacher who made the course on Esperanto has evaluated these productions.
- The answers to a questionnaire about their opinion on the activity and the proposed environment. 15 participants answered the questionnaire at the end of the study.

6 Results

6.1 General Results on the Activity

We first report the observations made by the teacher who monitored the work of the participants during the study:

- The members of the team A (control group) allocated tasks and performed only the requested work for the activity.
- Only two members of the team B (control group) were involved in the project, the other two members have encountered technical problems.
- The team C (test group) has been the most active and has produced most advanced game.
- The members of the team D (test group) have not worked at home; they did not give any production.

We analyze in this section the responses of the students to the questionnaire with a global approach. We also report their opinion on the modding activity they carried out: the interest for the modding activity, the learning opportunity and the interest of playing the game modded.

According to Table 1, almost all participants were interested in the collaborative modding activity. Furthermore, most of them enjoyed the activity. Those who did not enjoy are the members of the team D who gave up the activity. These results confirm the potential of this type of activity for the involvement of the learners. As stated by a participant: "*Although I'm not a game player, I liked a lot this way to learn Esperanto*"[1].

Table 1. Responses to the questionnaire on the interest of the activity (n=15)

	Not at all	Not much	Yes a little	Yes
"Were you interested in the activity?"	0%	7%	13%	80%
"Did you enjoy the activity?"	0%	13%	33%	53%

We asked the participants about the pleasant aspect of this educational approach:

[1] These responses in French have been translated into English in accordance with their original meaning.

- "*Finally, we feel we have created something (sense of being useful), as opposed to a project whose outcome is simply a PowerPoint presentation or a report. I find it is rewarding*".
- "*Creating a game that aims to teach Esperanto to someone else is a good idea because, firstly, we learn Esperanto content by creating this game and, secondly, we have the feeling that it could be useful for other learners!*".
- "*The fact that the learning activity leads to a concrete "object", in this case a game*".

As shown in Table 2, the results let us think that the activity helped participants to learn Esperanto. They not only learned during the course, but also by modding the game and by playing the game developed by the other participants. Those who did not learn Esperanto are the members of the team D who gave up the activity.

Table 2. Responses to the questionnaire on the way to learn Esperanto during the study (n=15)

	Not at all	Not much	Yes a little	Yes
"Did you learn notions of Esperanto by this activity?"	0%	7%	20%	73%
"Did you learn notions of Esperanto by trying to mod the game?"	13%	0%	60%	27%
"If you have tested the modifications made by the others, did you learn new vocabulary?"	13%	7%	33%	47%

We also asked the participants if the aim of the activity (to learn Esperanto in a ludic way) has been reached. The responses are encouraging (1):
- "*Yes, because I felt that I was playing so I had fun trying to mod the game. I realized once the experiment is finished that I remember very well what I have learned in a ludic manner*".
- "*Yes! For once this is a subtle way to tackle pedagogy, placing barriers to work a language in a clever manner, without thinking about it. A success in my opinion*".

We then asked the participants for comments about the interest of playing the game developed by other participants (1):
- "*It was a good idea, and we continue to learn when we try the other games*".
- "*This is very interesting, because it allows you to see other people's ideas and to test other Esperanto exercises*".

These comments show that modding can be a peer learning activity as learners can learn with the productions of the others.

6.2 Results on Modding with Integrated Collaboration Tools

In this section, we analyse the results per team, so as to study the impact of the use of integrated collaboration tools (test group) on the activity. We have first observed the games developed by the participants. We present in Table 3 the level of the productions of the four teams for the modding activity, according to two criteria:

- The level of modifications made in the game: simple (copy of rules by changing the text, that was sufficient to follow the instructions) or advanced (modifications of rules, advanced modifications of the behaviour of the game).
- The level of integration of elements of the domain (Esperanto): only vocabulary taught during the course; also concepts taught during the course; or new concepts or vocabulary on Esperanto.

Concerning the accessibility of the modding activity, 4 participants that had no programming skills succeeded in making advanced modifications in the game. More generally, the results show that previous programming skills have no impact on the level of the productions delivered by the students and on their involvement in the activity.

Half of the participants have tested the modifications made by the other members of their team: 3 members of the team B and 4 members of the C. So the availability of integrated collaboration tools had no obvious impact on this result, as there is no difference between the control group and the test group.

Table 3. Productions of the participants according to the programming skills (n=16).(P = Programming skills, M = Modifications in the game, U = Using vocabulary and concepts, T = Testing other games).

Group	Control group								Test group							
Name of the team	A				B				C				D			
Has programming skills		P	P					P	P				P		P	P
Made simple modifications	M	M	M	M	M	M			M	M	M	M				M
Made advanced modifications	M				M				M	M	M				M	
Used vocabulary taught during the course	U	U	U	U	U	U	U		U	U	U	U				U
Used concepts taught during the course	U				U		U		U	U	U	U				U
Used new concepts about Esperanto					U				U	U		U				
Tested the modifications made by the other members of his/her team		T			T	T	T		T	T	T	T				

While the results for the two control teams are rather similar, we observe that the results for the control teams are very different. While the team C has produced high-level deliverables, the team D gave up. As the members of the team D have gave up the activity, we cannot conclude on their way to carry out the activity. However, we observe that most of the members (3) of the team C have made advanced modifications in the game, against only one member in the teams A and B. Furthermore, all the 4 members of the group C have integrated concepts taught during the course, against only one in the group A and two in the group B. Furthermore, 3 members of the group C have integrated new concepts in the game, while only one people in the other groups (group B) done this high-level activity. As a complement to these observations, 3 members of the team C declare in the questionnaire that the SVN tool helped them to share the productions within their team (one non-response). We can so conclude that, in this exploratory study, the use of integrated collaboration tools helped the team C to produce more high-level productions than the teams A and B. We can also notice that programming experience have no influence in the acceptance, neither in the ability to deeply modify the game (4 advanced modifications on 6 have been done by learners without programming skills).

The collaborative tools were also significant for the teacher, as they helped to manage the group. The forum was a way to send a message to all the members of a team, and the SVN was an easy way to access to their work when it was over. For instance, the production of team C was available on the SVN when it was time to give it back, whereas the control group teams have sent the work in several parts, so the teacher had to merge them manually.

7 Conclusion

In this paper, we have presented collaborative modding as a new way to explore learning games for educational purposes. Collaboratively modding games not only help learners to learn the content of the games, but also to learn to collaborate. Based on these assumptions, we proposed a software architecture to support such an activity. This architecture has been implemented in an environment composed of four elements: a game, a Game Development Kit, a contextual forum and a version control software. The results of an exploratory study show that this approach is accessible to non-programmers, that it is of interest to learners and that they can learn the concepts of a domain (the Esperanto language in this case). This study also shows the need for collaborative tools integrated into the modding environment to help learners to make advanced changes in the game. We are aware that such an activity requires more time than more traditional ways of teaching for both teachers and learners. Thus a teaching activity should not be based only on collaborative learning by modding, but it should be a part of it.

Although this first exploratory study was conducted with a rather few participants, it leads us to identify the potential of such an environment. On a short term, we should focus our work on the development of an environment that better integrates all the needed functionalities on a unique interface. We should also integrate the knowledge map to test its utility in a following experiment. In fact, we believe that this innovative learning environment requires a larger scale experiment that could provide more quantitative results.

Furthermore, we have proposed several tools for students but rather few for the teachers and/or tutors. One future issue of this work is the development of an assistant system for the teachers to prepare the modding environment. This system would help them to build the knowledge map, to link the knowledge with the elements of the game and then to monitor the students during the modding activity and to assess their productions.

References

1. Ang, C.S., Zaphiris, P., Wilson, S.: Wiki-supported Collaborative Narrative Construction in Game Communities. In: ECSCW 2005 Workshop on Computer Games CSCW, Paris, France (2005)
2. Becker, K.: The Magic Bullet: A Tool for Assessing and Evaluating Learning Potential in Games. International Journal of Game-Based Learning 1(1), 19–31 (2011)
3. Cignoni, G.A.: Reporting about the Mod software process. In: Ambriola, V. (ed.) EWSPT 2001. LNCS, vol. 2077, pp. 242–245. Springer, Heidelberg (2001)
4. Djaouti, D., Alvarez, J., Jessel, J.P.: Can Gaming 2.0 help design Serious Games?: a comparative study. In: Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games, pp. 11–18 (2010)
5. Djaouti, D., Alvarez, J., Jessel, J.-P.: Classifying Serious Games: The G/P/S Model. In: Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches, pp. 118–136. IGI Global (2011)
6. El-Nasr, M.S., Smith, B.K.: Learning through game modding. Computers in Entertainment (CIE) 4(1), 1–20 (2006)
7. George, S.: Contextualizing Discussions in Distance Learning Systems. In: Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies, Joensuu, Finland, pp. 226–230 (2004)
8. Lave, J., Wenger, E.: Situated Learning. Legitimate Peripheral Participation. Cambridge University Press (1991)
9. Loh, C.S., Byun, J.H.: Modding Neverwinter Nights into serious games. In: Digital Simulations for Improving Education: Learning Through Artificial Teaching Environments, pp. 408–426 (2009)
10. McAtamney, H., O'Shea, B., Mtenzi, F.: Using the Crytek game engine in the Dublin Institute of Technology. In: Proceedings of the 7th International Conference on Computer Games, Angoulême, France (2005)
11. Monterrat, B., Lavoué, E., George, S.: Learning Game 2.0: Support for Game Modding as a Learning Activity. In: 6th European Conference on Games Based Learning, Cork, Ireland, pp. 340–347 (2012)
12. Moshirnia, A.: The educational potential of modified video games. Issues in Informing Science and Information Technology 4, 511–521 (2007)
13. Oblinger.: Games and Learning. Educase Quarterly 29 (3), 1–7 (2006)
14. Postigo, H.: Video Game Appropriation through Modifications: Attitudes Concerning Intellectual Property among Modders and Fans. Convergence: The International Journal of Research into New Media Technologies 14(1), 59–74 (2008)
15. Scacchi, W.: Modding as a Basis for Developing Game Systems. In: Proceeding of the 1st International Workshop on Games and Software Engineering, Waikiki, Honolulu, USA, pp. 5–8 (2011)
16. Tavares, J.P., Roque, L.: Games 2.0: Participatory Game Creation. In: Proceedings of the 6th Symposium on Computer Games and Digital Entertainment, São Leopoldo, Brazil (2007)
17. Volk, D.: Co-creative game development in a participatory Metaverse. In: Proceedings of the Tenth Anniversary Conference on Participatory Design 2008, Bloomington, IN, USA, pp. 262–265 (2008)

When One Textbook Is Not Enough: Linking Multiple Textbooks Using Probabilistic Topic Models

Julio Guerra[1,2], Sergey Sosnovsky[3], and Peter Brusilovsky[1]

[1] University of Pittsburgh, 135 North Bellefield Ave., Pittsburgh, PA. 15260, USA
{jdg60,peterb}@pitt.edu
[2] Universidad Austral de Chile, Independencia 641,Valdivia, Chile
[3] CeLTech, DFKI, Campus D3.2, D-66123 Saarbrücken, Germany
sosnovsky@gmail.com

Abstract. The Web-revolution in publishing and reading is rapidly increasing the volume of online textbooks. Nowadays, for most of the subjects, a selection of online textbooks is available. Such an abundance leads to an interesting opportunity: if a student does not like how a primary textbook presents a particular topic s/he can always access its alternative (e.g. more detailed or advanced) presentation elsewhere. Modern e-learning environments could better support access to different versions of instructional material by generating intelligent links between the textbooks sections that present similar topics and concepts. This paper reports an attempt to investigate the problem of fine-grained intelligent linking of online textbooks based on the probabilistic topic modeling technology. Using collections of textbooks in two domains (Elementary Algebra and Information Retrieval), we have demonstrated that intelligent linking based on probabilistic topic models produces a much better modeling quality than traditional term-based approaches.

Keywords: Hypermedia, Textbooks, LDA, Topic Model, Document Linking.

1 Introduction

The vast amount of learning content available on the Web opens several opportunities for adaptive educational systems supporting learners in finding the "right content". Practically, in any domain, a learner can easily find online dozens, hundreds or even thousands of instructional texts, tutorials and electronic textbooks. However, open-corpus resources are heterogeneous, they may organize their content in very different ways: by focusing on different parts of domain knowledge and covering the same parts with different levels of details, by using different terms for the same concepts (synonymy) and the same terms for different concepts (polysemy), by aggregating and structuring domain knowledge with different sets (or hierarchies) of chapters, sections, pages, etc. [9]

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 125–138, 2013.

The abundance of online textbooks results in interesting opportunities for modern students: those not satisfied with presentation of a particular topic in their primary textbook have a chance to access an alternative presentation from a range of textbooks covering this topic. To turn this opportunity into practice, modern e-learning environments should support their users by generating dynamic links between sections of different books that present similar topics and concepts.

For example, imagine a learner reading a section about "Linear Equations in Two Variables" from an Algebra textbook. The learner has problems understanding some parts of this section and decides to search for other textbooks covering similar content in her/his e-learning system. Since the system has access to multiple educational resources in this domain and understands semantic relations between their parts and sections, it returned a ranked list of links to relevant content of the other textbooks: a chapter titled "Linear equations (part II)" in one book, a section titled "Solving linear systems of equations" in another book, a subsection in the same book titled "Graphical solving of linear equations in two variables", and a chapter titled "Solving equations" in a third book.

The idea of automatic links generation between related fragments of educational content has been pioneered by Mayes and Kibby [7]. The keyword-level similarity-based approach suggested by them has been applied in a number of systems and architectures since[5]. Unfortunately, the practical quality of simple keyword-level similarity linking appeared to be less than perfect, as it can often link pages that were not really meaningfully related, which would result in presenting a wrong page to the user. While similar technologies have been often satisfactory in other domains, low quality of linking is unacceptable for e-Learning systems. As a category of users, students are very susceptible to the system's failures and are rarely able to recover from those on their own. Erroneous instructional material presented to a student can lead to confusion, frustration, lack of motivation and, essentially, poor learning [6].

Another area of research that has addressed the problem of linking documents based on their meaning is Semantic Annotation. Several systems have been built capable of enriching online content with dynamic links to relevant documents (e.g. COHSE [1] and Magpie [4]). Such systems are very good in recognizing special named entities (such as names, places, organizations as well as quantitates, dates, times etc.), and detecting keywords that match concepts names in a source ontology or a thesaurus; unfortunately, they disregard the rest of the content. This is an effective approach for general information access, but it is ill-suited for the educational domain. Instructional content rarely contains specially formatted entities; and one cannot expect that similar documents in two different textbooks from an arbitrary subject will be always using the same keywords.

This paper reports an attempt to re-investigate the problem of fine-grained intelligent linking of online textbooks using a popular approach to probabilistic topic modeling known as Latent Dirichlet Allocation (LDA). LDA [2] is a statistical model that automatically extracts topics from a collection of documents. Over the last few years, this technique has been applied successfully for discovering semantic models in large, heterogeneous and unstructured (or lightly-structured) collections, such as scientific journal papers or collections of news posts [2,3]. However, there is no known attempt of using this technique for modeling hierarchical structured collections, such as textbooks for the task of document linking. Our challenge has been to explore whether LDA can be applied within domain-specific collections of hierarchically structured educational content and can successfully support intelligent linking of online textbooks.

A study presented in this paper uses a collection of textbooks in two domains (Elementary Algebra and Information Retrieval) to explore whether intelligent linking based on probabilistic topic models can achieve a better quality of section-level textbook linking than previously used term-based approaches. To maximize the quality of the new technology, we also explore some important parameters associated with the application of LDA-based approach in hierarchical textbook context and report the performance results.

2 Linking Documents Using LDA

LDA topic models are generative probabilistic models that conceive each document as a mixture of topics and each topic as a mixture of words. LDA represents documents by a probability distribution over the topic space and defines topics as a probability distribution over the vocabulary of the collection. For building the model, a collection of documents is inputed to LDA. The algorithm iteratively assigns a topic to each word in each document and adjust the topic-word probabilities. The process can be seen as an optimization problem, in which the LDA mechanism continuously minimizes the number of "highly probable" topics per document and, at the same time, minimizes the number of "highly-probable" words per topic [2,10]. This leads to topics incorporating words that often occur together. LDA receives as inputs the number of topics to discover and the two prior (hyper) parameters to control both the sparsity of topic distributions in the documents and the sparsity of word distributions in the topics (the LDA setup for this specific project is explained in more details in Section 4.3). Once the model has been built, each document is represented as a probability vector over all the topics, and each topic as a probability vector over all the words in the vocabulary. Documents are discriminated based on different concentrations of topic probabilities. Taken the topic representations of the documents, the homogeneity between them can be computed using distance or similarity measures [10]. In our case, once we have a vector representation of all chapters, sections and subsections of the textbooks, we can proceed to compute similarity among parts of different books and, for each book part, create a ranking of similar other-books' parts. As we will see in the next section, there are some issues regarding the hierarchical structure of the textbooks to consider in order to obtain meaningful topic distributions of all book parts.

3 Research Questions

The main goal of this work is to explore the use of probabilistic topic models for generating high-accuracy linking of textbook parts (chapters, sections and subsections). The LDA-produced topic models are examined against the standard term-based models. The first research question is:

Q1 *Will probabilistic topic modeling produce a more accurate linking of textbook parts than the common term-based approach?*

The key aspects to consider are the characteristics of textbooks. Textbooks are hierarchically organized and the content of each chapter, section or subsection is the aggregation of its children's contents. The assignment of topic distribution among parts of the textbooks on different levels should consider these characteristics for a correct representation of each document's content. This leads to the second research question:

Q2 *How to incorporate hierarchical structures of textbooks in probabilistic topic models?*

We need a topic model to correctly represent each document (each textbook part) as the aggregation of the content of its sub-parts. One option is to build the topic model from all textbook parts with aggregated content, for example, a section aggregates the content of the subsections and in this form is input to LDA. However, this approach seems to "confuse" LDA as demonstrated by results of our preliminary tests. Other, more reasonable approach is to consider building the model using only documents that have actual (textual) content (usually terminal or leaf nodes, i.e. sub-sections) and further options for indexing the intermediate nodes (i.e computing the topic distribution of chapters and sections). We consider two approaches:

a) Aggregate topic distributions along the hierarchical structure of the textbook while weighting sub-topics' distributions by their sizes. For example, the topic distribution of a section is the aggregation of the topic distributions of it's subsections. The following formula explains the aggregation of topic distributions for a specific book part d.

$$\Psi'_d = \frac{\sum_{c \in C_d} \left(s_c \Psi_c\right) + s_d \Psi_d}{\sum_{c \in C_d} \left(s_c\right) + s_d} \qquad (1)$$

Ψ_d and Ψ'_d are the topic distributions of the node d before and after the aggregation, respectively; C_d is the set of child nodes of d, Ψ_c is the topic distribution of the child c, s_c and s_d are the sizes measured as the number of words in the respective document. For simplicity, hereafter, we refer to this option as *Topic Aggregation* (TA).

b) Re-index the aggregated content of intermediate documents using the inference mechanism provided by the topic model. This means, after the model is built, a version of the collection is created, in which each chapter and section contains the text content of all its children nodes. Then, each aggregated document is inputed to the built LDA model to obtain its new topic distribution. From now on, we refer to this approach as document *Re-Indexing* (RI).

Additionally, since individual textbook models reflect corresponding individual views on the domain stemming from differences in terminology that different authors may use, we are interested to explore the potential added value of building unified topics model using several books. We expect that a model built using multiple textbooks will reflect a better (more complete and objective) understanding of the domain and will support better document linking than a model built using a single textbook. For simplicity, we further call these options *Multiple Book* (MB) and *Single Book* (SB), respectively. Thus, the third research question is:

Q3 *Will the model built using multiple textbooks (MB) produce a better document linking compared to a model built using a single textbook (SB)?*

4 Experiment Design

We have conducted several experiments building LDA from textbooks in two domains: Elementary Algebra and Information Retrieval. Resulting topic models are used for computing similarity between the parts of different books in each domain. The evaluation approach compares the list of similar documents found by the LDA models and the baseline term-based model with an "ideal" document mapping provided by experts.

4.1 Textbooks

Four Algebra and five Information Retrieval textbooks were used in the study. For further references, the textbooks are labeled as BOOK1, BOOK2, etc. For the both domains, topic models in the Single Book condition (SB) are built based on BOOK1, and BOOK2 is used for evaluation. For Algebra, BOOKs 1, 3 and 4 are used for building topic models in the Multiple Book condition (MB); and, in the Information Retrieval domain, BOOK5 is additionally used for this. A sequence of technical procedures was performed on the contents of the textbooks: the text was converted to lowercase, stop-words and additional frequent words in the domain (i.e.: "exercises", "solutions" in Algebra) were removed.

Elementary Algebra Textbooks.

1: Elementary Algebra, by W. Ellis & D. Burzynski.
2: Elementary Algebra - v1, by J. Redden.
3: Understanding Algebra, by J. Brennan.
4: Fundamentals of Mathematics, edited by D. Burzynski & W. Ellis.

Information Retrieval Textbooks.

1: Introduction to Information Retrieval, by C. Manning, P. Raghavan & H. Schütze.
2: Modern Information Retrieval, by R. Baeza-Yates and B. Ribeiro-Neto.
3: Finding Out About, by R. Belew.
4: Information Storage and Retrieval Systems, by G. Kowalski.
5: Information Retrieval, by C. van Rijsbergen.

4.2 Ground Truth

As the ground truth, we used manual mappings of the two textbooks (BOOK1 and BOOK2) made by groups of experts in both domains. Overall, ten experts contributed: one professor and six PhD students from the School of Information Sciences at the University of Pittsburgh; three researchers from the Centre for e-Learning Technologies at DFKI. In the Algebra domain, five chapters of BOOK1 were mapped to BOOK2. In the Information Retrieval domain, four chapters of BOOK1 were mapped to BOOK2. To obtain more objective judgements, for each chapter, two experts were providing the mapping. A dedicated Web-interface was developed to facilitate the mapping task. Specific instructions were given in order to have a consistent mapping results: i) every chapter, section and subsection of BOOK1 had to be mapped to zero or more parts from BOOK2; ii) mapping had to be as accurate as possible; iii) mapping could relate book parts from different levels (for example, sections could be mapped to chapters, sections, or subsections); iv) it had to be taken into account that the content of a textbook part is the aggregation of the content of its subparts. Additionally, experts had to assign to every mapping a level of relevance and a level of confidence, both ranging from 1 to 3 (low, medium, high). A score for each match was computed by multiplying the relevance and confidence levels. The final score of each match was computed as the sum of the scores provided by both experts. Non-matched parts of BOOK2 were assigned with the zero score. Finally, the ground truth was represented as the compilation of all mappings blended into a single list: each element is a part of the BOOK1 and the respective ranked list of all matches from BOOK2 with their final scores.

4.3 LDA Setup

We used implementation of LDA provided by the MALLET Toolkit [8]. In MALLET, LDA setup depends on the number of topics, the number of sampling iterations, the smoothing over document-topic distribution hyper-parameter α, and the smoothing over topic-word distribution hyper-parameter β (a good explanation of the LDA hyper-parameters can be found in [10]). We set the number of iterations to 2000, taking into account the size of the documents and the collections. For selecting the number of topics, we followed a simple approach: since we expect the topics to represent semantic units of textbook content, we estimated that the number of topics should be between the number of sections and the number of subsections in an average book (sometimes, sections covers several topics, and, sometimes, subsections cover examples, exercises or different views of the same topic). In the Algebra domain, BOOK1 has 74 sections and 228 subsections; thus, we chose the number of topics equal to 150. This number also gave us the best results in the preliminary tests. In the Information Retrieval domain, the BOOK1 has 120 sections and 178 terminal nodes (some sections do not have children and are counted also as subsections). Here we also chose 150 as the number of topics. With regards to LDA hyper-parameters, we set up initial values of $\alpha = 0.01$ and $\beta = 0.01$, and then used the fixed-point optimization for hyper-parameters [11] implemented in MALLET.

4.4 Measuring the Effectiveness of Document Linking

We evaluate the effectiveness of all compared models on the task of finding documents similar to a target document using average NDCG@1, @3 and @10[1] as follows:

1. for every part of BOOK1, each part of BOOK2 is ranked according to the similarity of their topic distributions. As a similarity measure, we used the reciprocal symmetric KL-divergence[2];
2. for every part of BOOK1, the ranked lists generated by evaluated models are compared against the respective rank list from the ground truth using NDCG@1, @3 and @10;
3. for every LDA condition, the average NDCG@1, @3 and @10 are computed over all parts of BOOK1;
4. the LDA sampling process uses random seeds and produces different topic sets on each run; for every LDA condition, we run the model 30 times, thus, obtaining N=30 data points.

Baseline. As the baseline we use the effectiveness (measured by NDCG@1, @3 and @10) of the ranked lists obtained as a result of querying an index built using Apache Lucene[3]. Lucene computes similarity between a query and the documents in the collection using the standard TF-IDF model. For each part of the BOOK1, we toss a query to the index built from BOOK2. The query returns a ranked list of the parts in the BOOK2, which is used to compute the baseline average values for NDCG@1, @3 and @10.

[1] NDCG (normalized discounted cumulative gain) measures the quality of ranking documents by relevance. It compares the target ranking to the positions that documents occupy in the ideal list and penalizes the mismatches. NDCG@1 measures the effectiveness of the model to find the top relevant document. In the same way, NDCG@3 and NDCG@10 measure the quality of ranking the first three and ten items, respectively.

[2] Kullback-Leibler (KL-) divergence computes the difference between two probability distributions. The reciprocal symmetric KL-divergence is a modification of the original formula that can be used as a similarity measure [10].

[3] `http://lucene.apache.org`

4.5 Conditions and Hypotheses

Q1 examines the idea of contrasting LDA-based models with the baseline. Q2 is focused on comparing two different strategies helping to incorporating textbook hierarchies in the topic model: *topic aggregation* (TA), and *re-indexing* (RI). We further refer to this as the *aggregation* factor. Q3 stresses the comparison between a model built using a *single book* (SB) with a model built using *multiple books* (MB) from the collection. We further refer to these options as the *books* factor. Thus, four conditions are defined as can be seen in Table 1. KL-Divergence is used as the similarity measure; it helps to produce ranked lists of links from the nodes of BOOK1 to the documents of BOOK2. Ranked list are evaluated against the ground truth using averaged values of NDCG@1, @3 and @10. Finally, the LDA models are sampled 30 times to produce 30 data points for every condition. The experiment is executed separately in both domains.

Table 1. The 4 model conditions

		Aggregation	
		Topic Aggregation (TA)	Re-Indexing (RI)
Books	Single Book (SB)	SB-TA	SB-RI
	Multiple Books (MB)	MB-TA	MB-RI

To address Q1, the four conditions are compared to the baseline values of NDCG@1, @3 and @10. Here, we define the following hypothesis:

H1 *At least one of the models built using LDA performs better document linking than the baseline.*

Since the baseline accuracy score is a fixed value, we use one-sample t-tests to verify H1. To answer Q2 and Q3, we state the following hypotheses:

H2 *The LDA model aggregating topic probabilities based on the book hierarchy (TA) will perform better document linking than the model using re-indexing of content-aggregated intermediate documents (RI).*

H3 *The LDA model built using several textbooks (MB) will perform better document linking than the model built using a single textbook (SB).*

For testing H2 and H3, the four groups are compared based on the 2x2 between-subjects ANOVA design (see Table 1). Interactions, main effect and marginal means are reported.

5 Results

5.1 LDA vs. Baseline

Mean values and standard deviations of NDCG levels for all LDA conditions together with the results of t-tests comparing them against the baseline are reported in Table 2. Figure 1 presents these data graphically.

Table 2. Effectivity of the four conditions in the Algebra domain (top) and Information Retrieval domain (bottom)

Algebra

	NDCG@1			NDCG@3			NDCG@10		
Baseline	*.3662*			*.5807*			*.6582*		
	Mean	Std. Dev.	Sig. (p)	Mean	Std. Dev.	Sig. (p)	Mean	Std. Dev.	Sig. (p)
SB-TA	.547	.025	$<.001$	.647	.018	$<.001$	.691	.015	$<.001$
MB-TA	.532	.036	$<.001$	.620	.021	$<.001$	.663	.017	.165
SB-RI	.456	.027	$<.001$	.601	.027	$<.001$	.675	.019	$<.001$
MB-RI	.414	.040	$<.001$	.572	.032	.132	.647	.026	.022

Information Retrieval

	NDCG@1			NDCG@3			NDCG@10		
Baseline	*.057*			*.186*			*.258*		
	Mean	Std. Dev.	Sig. (p)	Mean	Std. Dev.	Sig. (p)	Mean	Std. Dev.	Sig. (p)
SB-TA	.345	.051	$<.001$	.461	.042	$<.001$	.536	.033	$<.001$
MB-TA	.309	.063	$<.001$	.418	.045	$<.001$	.520	.039	$<.001$
SB-RI	.360	.066	$<.001$	.484	.053	$<.001$	.556	.045	$<.001$
MB-RI	.336	.050	$<.001$	.456	.054	$<.001$	.534	.041	$<.001$

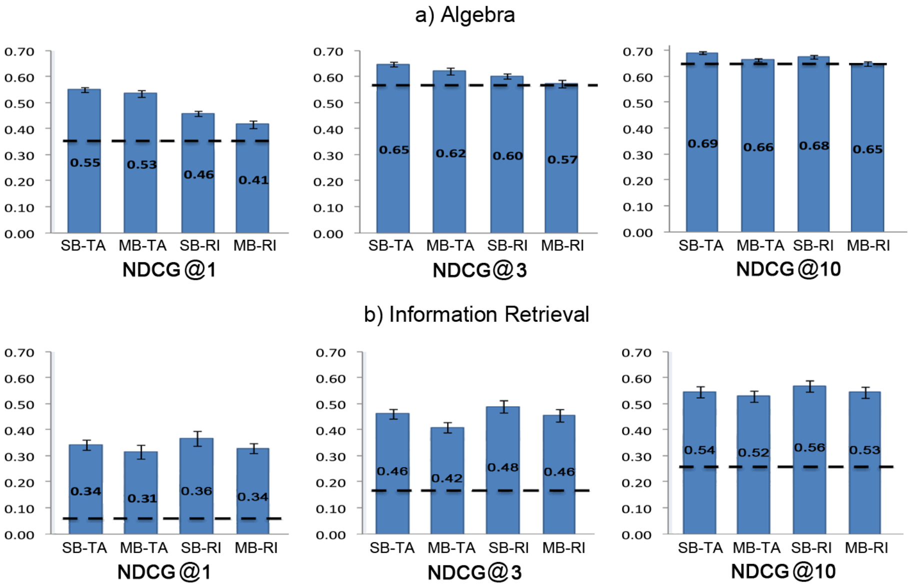


Fig. 1. Four conditions individually compared with the baseline in Algebra and Information Retrieval domains. Baseline NDCG@1,@3 and @10 values are indicated with dotted lines.

Algebra Domain. As it is seen from Table 2, for Algebra domain, all LDA models for all NDCG levels produce significantly better results than the baseline except for NDCG@3 of MB-RI, NDCG@10 of MB-TA and NDCG@10 of MB-RI (which is even significantly lower than the respective baseline value). Among the significantly better LDA scores, the highest effect sizes were always presented by the SB-TA model outperforming the baseline by several standard deviations (NDCG@1: $p < .001$, Cohen's d = 7.232; NDCG@3: $p < .001$, Cohen's d = 3.683; NDCG@10: $p < .001$, Cohen's d = 2.187). Since for all NDCG levels there was at least one condition that performs significantly better than the baseline, these results support H1 in the Algebra domain.

Information Retrieval Domain. Within the Information Retrieval collection, all LDA conditions perform significantly better than the baseline (see the bottom part in Table 2).

H1 is Supported. Overall, LDA produces significantly better document linking than the baseline term-based model in both domains. The highest results are always obtained for NDCG@1, which means that the LDA-based approach performs much better in finding the best matches, i.e the most similar documents. It is also very interesting to note the difference between the two domains. Information Retrieval is a newer domain than Algebra and, thus, has a less standardized vocabulary. Therefore, LDA is especially effective compared to the standard term-based approach. This result confirms the power of probabilistic topic modeling in domains with sparse and non-standardized vocabularies.

5.2 TA vs. RI

For testing H2, in each domain, a 2x2 between-subjects ANOVA was performed on the scores of NDCG@1, NDCG@3, and NDCG@10 as a function of *aggregation* strategy (TA, RI) and *books* strategy (SB, MB). Table 3 presents the marginal means and standard error values for all conditions and domains.

Table 3. Marginal Means

		NDCG@1		NDCG@3)		NDCG@10	
Algebra		M	SE	M	SE	M	SE
Aggregation	TA	.539	.004	.633	.003	.677	.003
	RI	.435	.004	.587	.003	.661	.003
Books	SB	.502	.004	.624	.003	.683	.003
	MB	.473	.004	.596	.003	.655	.003
Information Retrieval		M	SE	M	SE	M	SE
Aggregation	TA	.327	.007	.439	.006	.528	.005
	RI	.348	.007	.470	.006	.545	.005
Books	SB	.352	.007	.473	.006	.546	.005
	MB	.323	.007	.437	.006	.527	.005

Algebra Domain. All ANOVA assumptions have been satisfied for all conditions and all NDCG levels.

The patterns of differences among *aggregation* strategies are significantly different between SB and MB only for NDCG@1, F(1,116) = 5.274, p = .023, $\eta_p^2 =$ .043. No other significant interactions have been observed. For the main effect of *aggregation*, the results show that TA produces significantly higher scores than RI for all NDCG levels averaged across *books* strategy (NDCG@1: F(1,116) = 309.496, p < .001, $\eta_p^2 = .727$; NDCG@3: F(1,116) = 103.181, p < .001, $\eta_p^2 =$ .471; NDCG@10: F(1,116) = 19.332, p < .001, $\eta_p^2 = .143$). These results support H2 in the Algebra domain. Additionally, it can be seen that the effect decreases with every next NDCG level. This underlines the important of choosing the right *aggregation* strategy when the goal is to find the most relevant documents from another textbook.

In order to find the pattern of differences in NDCG@1 scores among *books* and *aggregation* strategies, a simple main effect of *aggregation* has been performed for SB and MB strategies. When LDA models are build using a single book, TA (M=.547, SE=.006) is significantly better than RI (M=.456, SE=.006): F(1,116) = 116.983, p < .001, $\eta_p^2 = .502$. When multiple books are used, TA (M=.532, SE=.006) also significantly outperforms RI (M=.414, SE=.006): F(1,116) = 197.787, p < .001, $\eta_p^2 = .630$. These results showed that the interaction between *aggregation* and *books* factors for NDCG@1 does not change the fact that TA works better than RI in the Algebra domain; it only means that, when using TA, there is no difference between SB and MB conditions.

Information Retrieval Domain. All ANOVA assumptions have been satisfied for all conditions and all NDCG levels, except for a slight deviation from normality of NDCG@3 under MB-RI condition (Shapiro-Wilk = .929, p = .047).

There is no significant difference in the patterns of differences among *aggregation* strategies between *book* strategies for all NDCG levels. The main effect of *aggregation* has been observed for all NDCG levels. However, it is the opposite of what we have seen in the Algebra domain. For Information Retrieval, TA produces significantly lower scores than RI (NDCG@1: F(1,116) = 4.017, p = .047, $\eta_p^2 = .033$; NDCG@3: F(1,116) = 11.903, p = .001, $\eta_p^2 = .093$; NDCG@10: F(1,116) = 5.667, p = .019, $\eta_p^2 = .047$). These results do not support H2 in the Information Retrieval domain.

H2 is Partially supported. In the Algebra domain, Topic Aggregation strategy performs better than Re-Indexing. However, the results are opposite in the Information Retrieval domain, where RI performs better than TA. This is interesting, as domain differences seem to matter when choosing the best *aggregation* strategy. One possible explanation can be inferred from the different characteristics of the two domains and the differences between the used textbooks. For example, the vocabularies that LDA models are processing in these domains are very different in size. In Algebra, a single book produces a vocabulary of 2,220 terms, while in Information Retrieval, a single book model works with the vocabulary of 13,405 terms. Further research is needed in order to better understand the differences among the domains and how they influence the topic modeling process.

5.3 MB vs. SB

In this section, we compare the performance of LDA models built from a Single Book (SB) with models built from Multiple Books (MB) in order to test H3. The ANOVA design described in previous sections is also used here as well.

Algebra Domain. In the Algebra domain, there are no significant interactions among *books* and *aggregation* factors except for NDCG@1. The simple main effect test reported in the previous section has shown that it does not matter, whether the LDA model is built from a single book or multiple books, when it uses TA as the *aggregation* strategy. For the main effect of *books*, the results show that, in the Algebra domain, SB produces significantly higher scores than MB averaged across *aggregation* strategies for all NDCG levels (NDCG@1: $F(1,116) = 24.180$, $p < .001$, $\eta_p^2 = .172$; NDCG@3: $F(1,116) = 38.361$, $p < .001$, $\eta_p^2 = .249$; NDCG@10: $F(1,116) = 62.095$, $p < .001$, $\eta_p^2 = .349$). Thus, H3 is not supported in the Algebra domain.

Information Retrieval Domain. Similar results are obtained in the Information Retrieval domain. There is significant difference between SB and MB strategies averaged across *aggregation* for all NDCG levels (NDCG@1: $F(1,116) = 7.849$, $p = .006$, $\eta_p^2 = .063$; NDCG@3: $F(1,116) = 16.099$, $p < .001$, $\eta_p^2 = .122$; NDCG@10: $F(1,116) = 6.871$, $p = .010$, $\eta_p^2 = .056$). All these effects strongly favor using the Single Book strategy over Multiple Books and do not support H3 in Information Retrieval domain as well.

H3 is Not Supported. In both, Algebra and Information Retrieval domains, the results have shown that LDA models built using a Single Book (SB) will produce significantly better document linking than the models built from Multiple Books (MB) regardless of the hierarchical content *aggregation* method.

6 Conclusions

In this work, we have explored the use of LDA-based topic models within collections of textbooks for the task of fine-grained cross-collection document linking. We have shown that LDA is a valuable alternative to the standard term-based approach and outperforms it, especially, for finding the most similar documents (NDCG@1, NDCG@3) and in less standardized domains. We have applied two different approaches for building topic models with regards to the hierarchical structure of textbooks and discovered that, in different domains (and, perhaps, content collections), different aggregation strategies can be better. We have also observed that the topic models built using one textbook produce a better document linking than the model built using multiple books, contrary to our expectations. Further research is needed to improve our understanding of these differences and their impact on the LDA model building.

We believe, that using LDA is a promising approach for addressing the problem of automatic and fine-grained textbook linking and it can be further applied to facilitate content modeling and adaptation in the open corpus settings. The ability of probabilistic topic-based models to help finding the top similar documents is a clear advantage over the traditional term-based methods and can be used to implement effective recommendation and navigation support in adaptive educational hypermedia systems. We plan to incorporate this technology in an e-learning environment, as well as to keep investigating mechanisms for improving the quality of LDA models. Our future work will include experimenting with the techniques for topic models evaluation, utilizing topics and textbooks structures to discover semantic relations among the topics, ensembling topic models with the models based on keyword and concept extraction techniques, and further investigating the application of this approach in other domains.

Acknowledgements. Julio Guerra is supported by Chilean Scholarship (Becas Chile) from the National Commission for Science Research and Technology (CONICYT, Chile), and the Universidad Austral de Chile.

References

1. Bechhofer, S., Goble, C., Carr, L., Kampa, S., Hall, W., De Roure, D.: COHSE: Conceptual Open Hypermedia Service. In: Handschuh, S., Staab, S. (eds.) Annotation for the Semantic Web, pp. 193–211. IOS Press (2003)
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research 3, 993–1022 (2003)
3. Blei, D.M., Lafferty, J.D.: Dynamic topic models. In: Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, June 25-29, pp. 113–120 (2006)
4. Domingue, J., Dzbor, M., Motta, E.: Magpie: supporting browsing and navigation on the semantic web. In: Rich, C., Nunes, N.J. (eds.) Proceedings of 9th International Conference on Intelligent User Interfaces (IUI 2004), Funchal, Madeira, Portugal, pp. 191–197. ACM Press, New York (2004)
5. Fountain, A., Hall, W., Heath, I., Davis, H.: MICROCOSM: An open model for hypermedia with dynamic linking. In: Streitz, N., Rizk, A., André, J. (eds.) Hypertext: Concepts, Systems and Applications, pp. 298–311. Cambridge University Press (1992)
6. Jednoralski, D., Melis, E., Sosnovsky, S., Ullrich, C.: Gap detection in web-based adaptive educational systems. In: Luo, X., Spaniol, M., Wang, L., Li, Q., Nejdl, W., Zhang, W. (eds.) ICWL 2010. LNCS, vol. 6483, pp. 111–120. Springer, Heidelberg (2010)
7. Mayes, J.T., Kibby, M.R., Watson, H.: StrathTutor: The development and evaluation of a learning-by-browsing on the Macintosh. Computers and Education 12(1), 221–229 (1988)

8. McCallum, A.: MALLET: A Machine Learning for Language Toolkit (2002), http://mallet.cs.umass.edu
9. Sosnovsky, S., Hsiao, I.-H., Brusilovsky, P.: Adaptation "in the Wild": Ontology-Based Personalization of Open-Corpus Learning Material. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 425–431. Springer, Heidelberg (2012)
10. Steyvers, M., Griffiths, T.: Probabilistic topic models. In: Landauer, T., McNamara, D.S., Dennis, S., Kintsch, W. (eds.) Handbook of Latent Semantic Analysis. Lawrence Erlbaum (2007)
11. Wallach, H.M.: Structured Topic Models for Language (Doctoral dissertation). University of Cambridge, UK (2008)

Research on Collaborative Planning and Reflection – Methods and Tools in the Metafora Project

Andreas Harrer[1], Kerstin Pfahler[1], Reuma de Groot[2], and Rotem Abdu[2]

[1] Catholic University Eichstätt-Ingolstadt, Ostenstr. 14, 85072 Eichstätt, Germany,
{andreas.harrer,kerstin.pfahler}@ku-eichstaett.de
[2] The Hebrew University of Jerusalem, 91905 Jerusalem, Israel
msruma@mscc.huji.ac.il, rotem_abdu@yahoo.com

Abstract. Collaboration in complex learning scenarios does not succeed automatically without structuring the learning process. In the literature there have been several means proposed to support collaborative learning: scaffolds, scripts, activity structures and reflection prompts are among these approaches. Scenarios that prompt and support reflection during the learning process proved to be effective for collaborative learning. In the Metafora project our approach is to support students in the process of learning to learn together (L2L2) by means of dedicated tools for planning and reflection. These tools allow students to organize their learning process in a self-regulated way. In this paper we will present the pedagogical background, the supporting tools and underlying design, and the research methodology used in Metafora to gain insights on how collaborative planning, learning, and reflecting interact. An extended example of qualitative analysis is given to illustrate our approach.

Keywords: Computer-supported collaborative learning (CSCL), learning to learn together (L2L2), self-regulated planning.

1 Introduction – Means to Support Learning

Collaboration in complex learning scenarios does not succeed automatically [1] without some means for supporting the learning process either by a computer system, a tutor/teacher, or peer help. In recent years several means have been discussed and investigated to support collaborative learning and reflection.

Creating structures for the learning process is an essential principle underlying several of these means: Both in activity structures [9] and in collaboration scripts [14] structure is imposed onto the learning process to help students in organizing their work. This can have the shape of external representations, such as a table for activity structures, to be filled in by the students; or it can be provided directly as instructions of desired behaviour and sequencing of activities as usually given in external collaboration scripts with the purpose that learners internalize these instructions for further use. Articulation and reflection are considered key elements for several instructional approaches to promote learning [11]. Reflection prompts [3] provided by a computer

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 139–150, 2013.

system to be filled in by the students to articulate and reflect on their thoughts, rationale, and actions are a concrete means to facilitate this cognitive activity.

Quintana et al. [11] bring these facets together in their scaffolding design framework. They give guidelines to support sense making, process management, and reflection in learning software. In the Metafora project we developed a learning system, with which school children of 12-16 years work together on complex scientific activities, using a multitude of learning tools in groups of 2-6 students. Scaffolding these learning process via a distinct organization space is relevant as well as the seamless transition between our diverse learning tools and the organization of artefacts created by the learners. Guidance for planning and monitoring of the students' own work and progress has been a rationale for the design of Metafora tools, especially the Planning Tool (cf. next section).

In our specific context of Metafora which is comprised of a set of diverse learning tools, we consider another concept of mediation across different "worlds": *Boundary objects* were coined by Star and Griesemer [13], who originally used them to "maximize both the autonomy and communication between worlds" in their study from 1989. They defined them as objects to depict the divergent use of information in divergent groups. Although they are plastic, a defining characteristic of boundary objects is that they contain enough changeless content to preserve a global identity. In 2004 Hoyles used boundary objects in a classroom contex: "a boundary object provides a generalized mechanism for meanings to be shared and constructed between communities" [6]. We will take up the concept of boundary objects as so called referable objects serving as mediating artefacts between different workspaces.

Based on these findings and on the project context of Metafora, where collaborative learning activities in complex scenarios using various tools are conducted, we came to the conclusion that empowering students to be aware of their own learning processes, in the context of self-regulated learning, can contribute to learning to learn together (L2L2). We will go into details in the next sections how planning & reflection support is conceptualized and implemented in the Metafora approach for this end.

2 The Metafora Approach to Support Learning to Learn together (L2L2)

In Metafora dedicated learning tools support the collaborative learning to learn together within groups. The tools comprise the Planning Tool, a place for self-regulated planning and reflection on learning activities, the LASAD argumentation environment to discuss activities [8], several microworlds [10] to interactively explore phenomena in the domains of math, physics and sustainability, and the educational game PIKI. All these tools are combined to experience complex learning scenarios.

Learning to learn together implies that all the group members are able to coordinate, regulate and plan the learning task by balancing issues of individual ability, motivation and expectation through constant dialogue. The process of L2L2 can be described and studied by analysing the groups' collaborative learning activities such as mutual engagement when performing collaborative tasks, distributed leadership, dialogue in appropriate spaces and formats, and peer assessment. How planning and reflection are supported is shown in the following.

2.1 Planning and Its Support

During our investigation into planning as a dedicated activity within the learning process, we identified and designed several principles and aspects that a specific Planning Tool for Metafora should have:

While project planning tools for commercial purposes stress mainly the temporal dimension in Gantt-like diagrams, we decided to allow a more flexible representation – called planning map - where students are able to use both dimensions of a computer screen to organize graphical icons representing their activities and connecting these icons with temporal and causal dependencies. Our design decision shares the principles of graphical modelling approaches for CSCL scripts as in the LDL modelling approach [4] or the CoSSICLE project [7] with the graphical MoCoLADe language [5]. In a literature research on science inquiry activities and argumentation processes the pedagogical partners in the Metafora project identified the most frequent elements of these learning processes. This resulted in the set of "planning cards" which is constantly updated and refined within the project. These cards contain titles and visual symbols that characterize various learning activities. These cards can be considered a visual language [15] with which the users describe their collaborative work. Because the number of cards of approx. 50 is too large to perceive at one glimpse, we categorized these cards according to their conceptual semantics and created the following categories:

- *Activity stages* as high-level activities, e.g. exploring a phenomenon
- *Activity processes* as methods and tasks of stages, e.g. discussing alternatives
- *Roles* for the organization and distribution of responsibilities, e.g. evaluator
- *Attitudes* as positions students can take up, such as being critical or creative
- *Resource cards* that serve as entry points to tools, e.g. a discussion workspace
- *Logical gates and cards for ad-hoc use*, such as a blank text card for annotations

Structuring the elements of the visual language according to the semantic categories reduces the load of structuring complex tasks for the students and thus can be considered following guideline 4 (provide structure for complex tasks and functionality) of Quintana et al. [11].

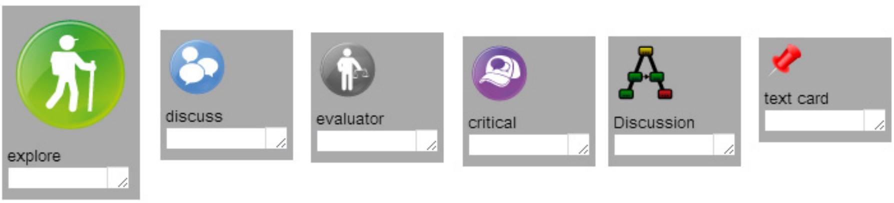


Fig. 1. One visual language element out of each of the different semantic categories from the Metafora Planning Tool to describe collaborative work, as listed above

As the planning space allows to follow-up on students' collaborative work it provides a natural flow and transition between planning actions and actions in the microworlds and in the LASAD discussion space. To achieve this, we implemented a flexible mechanism that allows associating planning cards with arbitrary web links that can be opened directly when choosing an appropriate option from a context menu for that card (see also resource cards below). This facilitates a *fluid transition between planning and enacting* as desired in our Metafora approach and in tune with guideline 6c (facilitate navigation among tools and activities) of the scaffolding framework [11]. While this works in principle for any web link (in our words “external tools”), the connection with Metafora tools is implemented with richer semantics: when opening a Metafora tool, the tool is customized with the user information, i.e. all tools are opened with the user context the current user is logged in with. Thus the user will get all the models, maps, and artefacts she/he has at her/his disposal in that tool. Moreover, at this point each individual of the group can also accesses materials created by other members of the group. Thus, users are able to review and collaboratively discuss the work that was done by the group. Technically this is implemented using auto-login procedures on the respective learning tools.

Another feature in the spirit of planning and enacting are resource cards, which are gateways to shared models and workspaces represented explicitly in the planning map. The general flow of using resource cards is:

- creating a resource card (e.g. “Discussion” card, Fig. 2, left) in the Planning Tool
- using it via a context menu (displayed in Fig. 2, left when option “start using” is chosen) on the card creates a workspace in the respective tool
- the link to this new workspace is deposited in the resource card and thus accessible at any later point in time.

When the user opens this card again, or other students working in the map open it, they will enter exactly the same workspace. Thus it is a means to share artefacts of the different microworlds via the planning map. In the example in figure 2, a user can create a new LASAD discussion map using a discussion resource card in the plan as described above and can make this map accessible to all other group members via the planning map. Resource cards can be considered as special boundary objects because of their mediating function.

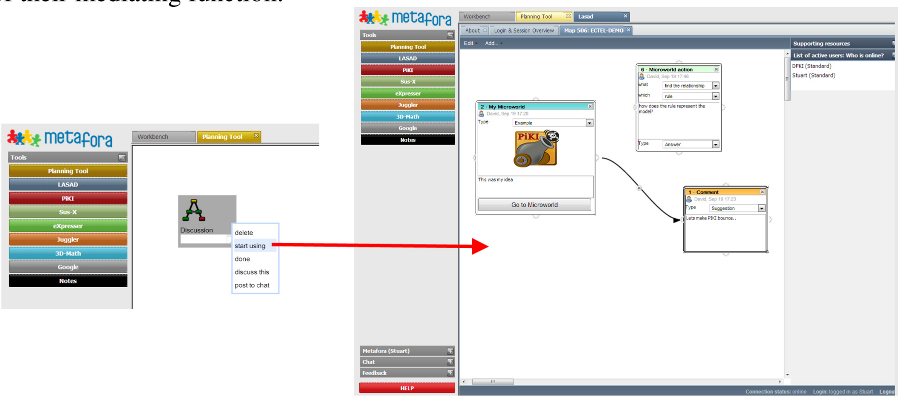


Fig. 2. A LASAD discussion workspace entered through a resource card in the Planning Tool

2.2 Reflection and Its Support

Besides pre-planning activities, the Planning Tool we designed is also meant to support the reflection on the students' learning process actively following guideline 7 of the scaffolding design framework of Quintana et al. [11]: "facilitate ongoing articulation and reflection during the investigation". This is especially important since in our early experiments we got indication that a full pre-planning is unlikely to happen, yet the continuous interplay of stepwise planning, reflection on the step, and further planning is present in most real activities conducted with the Metafora system.

To highlight the importance of joint representation of the current state of activities and the process as a whole we integrated functionality that allows the students to reflect and document what they have been doing in the process: in the visual state representation of each card there are three options, one for "planned, but not yet started" which is the default grey representation of a card, one for "planned and started" in yellow which is explicitly set when students start to use a card, and one for "planned and finished" in green which is set when students choose the option "done" in a context menu. The representation has been designed in analogy with checklists, where the marking of an item represents the fulfilment. One advantage we see in our representation is that the colour coding allows a one-glimpse perception of the overall process progress, expressed by the level of "greyness" or "greenness" a planning map has. Additionally, a closer look allows for checking each card specifically for the currently assigned status of work done on it. You can see an example showing the visual state representation for reflection and documentation in Figure 3 below.

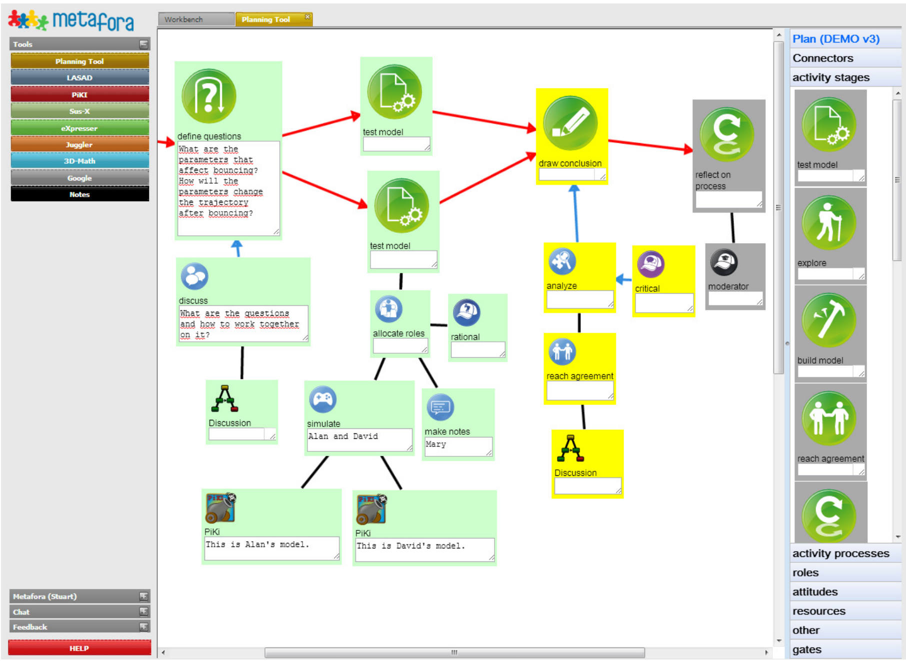


Fig. 3. Screenshot of the Planning Tool with process documentation – yellow cards have been started and are currently used, green cards have been finished by students' explicit marking

Another means for reflection about others' contributions and deictic reference [12] to artefacts produced is provided by us through referable objects that mediate between different tools and workspaces. Referable objects are not a new concept, but rather were inspired by the concept of boundary objects as discussed in the introduction. We found a need for boundary objects in Metafora because students need to share individual elements, or objects, from one tool with another tool to allow for reflection and discussion about the workings of one tool within another. When a group can share these objects, it is then possible to use the most appropriate tool to discuss and evaluate ideas focused on these objects, rather than being tied to the tool from which the object originates. This ability to visualize and discuss an individual object from one tool within another is powerful and helpful to students, but there is an obvious weakness: viewing an object in isolation can lose the context and meaning of the same object when located in its original environment. To alleviate this issue, and allow any user to see the full context of the given object, a reference is also recorded with the referable object, an address to the source of a tool where the object was created. Through this reference learners are able to go directly to the place where the element exists and can take a look at it. So we can see that our concept of referable objects expands upon the concept of a boundary object, and as such we offer the following definition of a referable object: "an individual element or product of any tool that is recorded by the system in such a way as to allow students, teachers and researchers to reference this object for discussion or evaluation at a later point in time." Referable objects can be shared in the system's group chat or can be made explicit in the discussion environment LASAD, such as the discussion object with a direct link to a saved game state of the ballistics microworld game PIKI (see the cannon object in figure 2).

3 The Combined Research and Pedagogy Approach in Metafora

For the evaluation in the Metafora project we chose a design based approach [2] for collecting data on L2L2 experiences and also on students' awareness of their L2L2 experiences and iteratively improve both system and pedagogy based on the (formative) evaluation results. As stated earlier, promoting groups' collaboration is not an easy task, especially when dealing with solving mathematics and physics problems, as the present project is focused on. Apart from the tool, that serves as an arena for collaborative L2L2 activities, the Metafora team designed an adequate pedagogy for the design of a Metafora-type learning activity. The "challenge"- an ill defined question that promotes various solution paths and sometimes even various solutions - is one of the pillars in any learning activity with the Metafora tool. After getting the challenge from their teachers, the students are asked to use the planning tool to plan their collaborative work, working in small groups of 2-4 students. While doing so, the students are also asked to refer to possible stages and processes of their work as explicated in the visual cards. From this point the group is expected to work on its own, moderated and assessed by their teachers or peers. Every challenge is expected to be solved during a certain period of time (~2 weeks). All the studies done by the pedagogical partners in Metafora were conducted in schools and in some cases were carried out by teachers in order to grasp the potentials and the drawbacks of using Metafora in the classroom. Considering the special conditions of each study - as it was influenced strongly by the partners' local conditions - the data collection was mainly based on videos and screen recording of the groups' work. In some cases partners conducted interviews or collected answers on specific questions related to students' attitudes and believe in collaborative group work. For data analysis we transcribed students' talk while they were working with the Metafora tools.

One of the studies that we show here as an illustrative case was taken with sixteen 8th grade students who met once a week for 1.5 hours in a framework of a mathematical club taken right after school hours, in the school. During this one year program the students worked collaboratively on solving mathematical problems and reflecting them. By the end of the school year students worked on two Metafora challenges.

The episode that we describe in the next section, took place while the class went through four iterations of solving a Metafora challenge in Math: the "City Challenge". In the challenge the students were asked to place an energy system in an imaginary city, in a point that is equidistant from seven other institutions. Since such a challenge is too complex to be solved in one step, the students had to solve it in four steps that were predefined by our research team (table 1). In addition to the Metafora tool the students use Geogebra, a Microworld that allows inquiry in Math for the creation of dynamic Geometric figures in a Cartesian domain (this tool is connected as an "external tool" to the Metafora platform).

Table 1. The design of the City Challenge in four distinctive steps

Step one: Place the energy centre in the middle of seven institutions, all of which are important to the city and all are packed with people. **Step two:** Simpler case: 3 institutions located in the structure of an equilateral triangle. Where to place the point? **Step three:** Three institutions located in the structure of any triangle. Where to place the point? Try to formulate a final conclusion. Is there an equally distant point in the three vertices of a triangle? If so where it is located? **Step four:** Give a general answer: Where should we put the energy centre?

In the next section we describe how a pair of two students used the Planning Tool for planning and reflecting in step three of the City challenge in a real classroom situation. The shown case covers only one facet of possible uses, both the remote usage and larger groups have been present in experimentation, too. This specific dyad was chosen because authors of this paper have been directly observing this dyad, making it an especially rich case for illustration. Generalised results and more quantitative evaluations are currently collected and will be published in forthcoming articles.

4 Example: Student dyad reflecting on the City Challenge

We follow a team of two students: S1 and S2. Both students already concluded in the previous session that a point that is equidistant from the three vertices of an equilateral triangle is the median's intersection point. They then tried to solve the problem posed to them in **step three**: locating the energy centre in the middle of any triangle. At this point the students were asked to report/reflect on what they did in steps 1 & 2 of the learning scenario, and plan their joint work for solving the challenge using the Planning Tool and Geogebra. In the following episode we present two observations made in the classroom when students use the Planning Tool.

- the first is about the way students elaborate their mutual understanding over scientific processes of inquiry while they use the visual language
- the second is about the affordances of the Planning Tool for students' interplay between their past, present and future activities.

In the second meeting of the challenge the students were asked to use the Planning Tool to reflect on what they did in the first meeting, where they used exclusively the Geometry tool as a preparatory activity. The reflection process was induced by the teacher to let the learners get familiar with the Planning Tool so that in later steps they will be able to plan in a self-regulated manner using the Planning Tool's intended main functionalities. Cards #1, #2 and #3 in fig. 5 exemplify this work. Figure 4 shows the oral discussion in its temporal flow that accompanied the work of S1 & S2 while they were collaborating in the Planning Tool from the same computer.

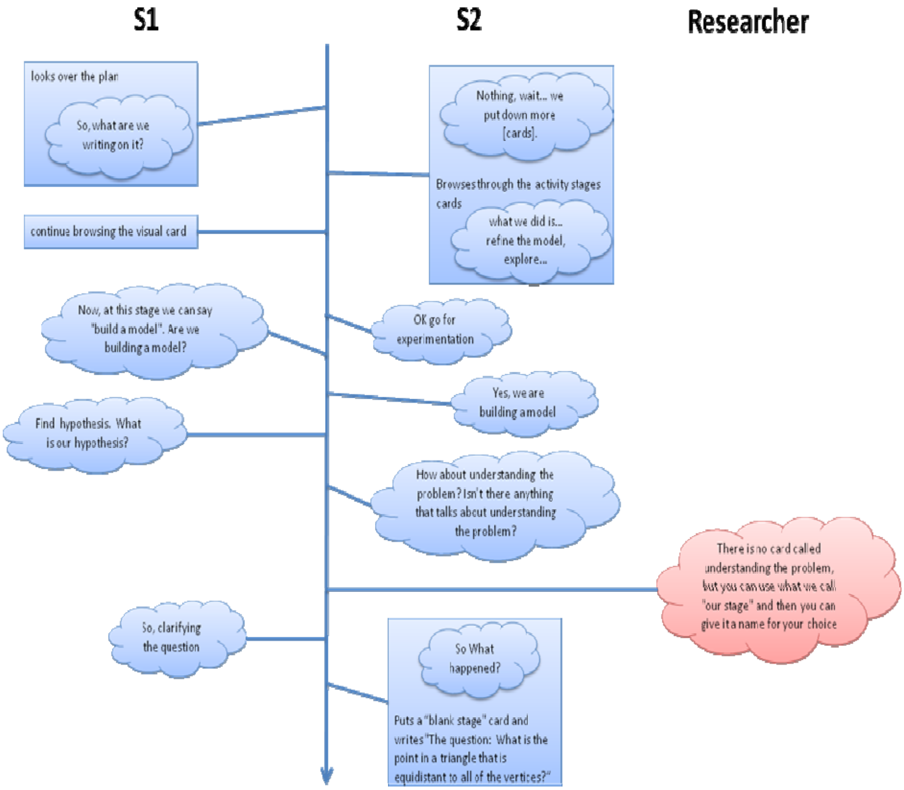


Fig. 4. Trace of the oral discussion between S1and S2 working from the same computer

The cards of the visual language give them ideas on how to explicate what they did (as can be seen in fig. 5). They start by putting card #1 "blank stage" in which they explicate what the problem is. They continue to card #2 "find hypothesis" with which they claim that they will try the same solution now as they tried in the case of equilateral triangle (and by that they imply their solution). In the case of card #3 "experiment" they attributed with the card comment "with Geogebra" to document the domain tool Geogebra they want to use for this purpose.

In the next episode S2 leaves and S1 keeps working alone. He starts by putting card #4 that presents the model in Geogebra that "will" be tested by them. Then he puts down an "intersection" card #5 that represents a decision intersection. If the model will "work"- they will draw conclusions (card #6 as a hypothetical stage- as written: "in case the model worked start drawing conclusion"). In case the model will not work they will have to come up with a new hypothesis and start another iteration of their inquiry cycle represented by card #7.

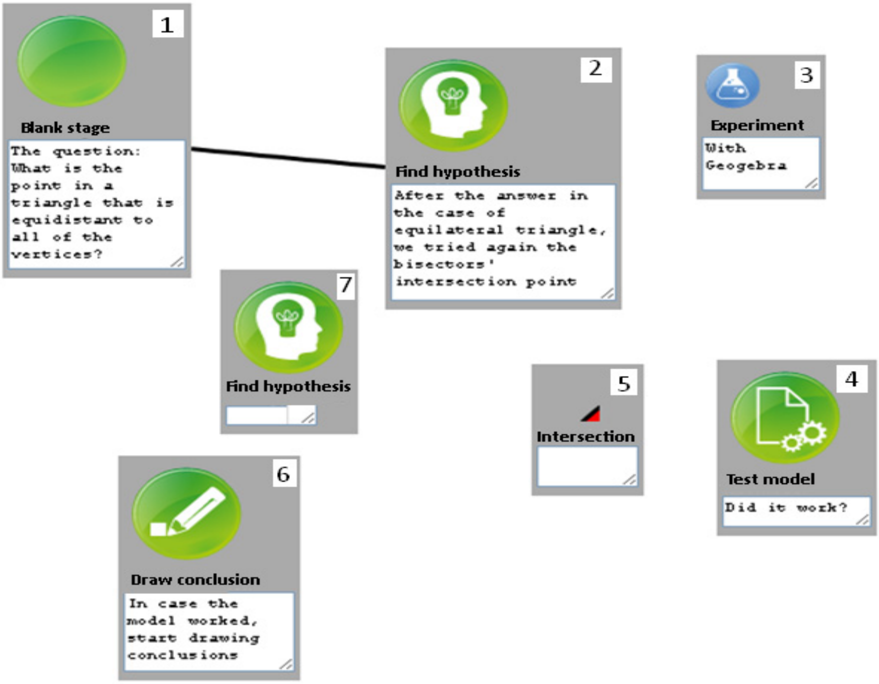


Fig. 5. The plan of students S1& S2 created with the Planning Tool

When S2 comes back and re-joins S1, he immediately takes the keyboard and starts writing in the "experiment" card #3 that the experiment they conducted did not work. S1 stops him and asks him to be patient and listen to his explanations. S2 listens and adopts the new concept / term "test model". The flow of oral discussion between the two students is shown in Figure 6.

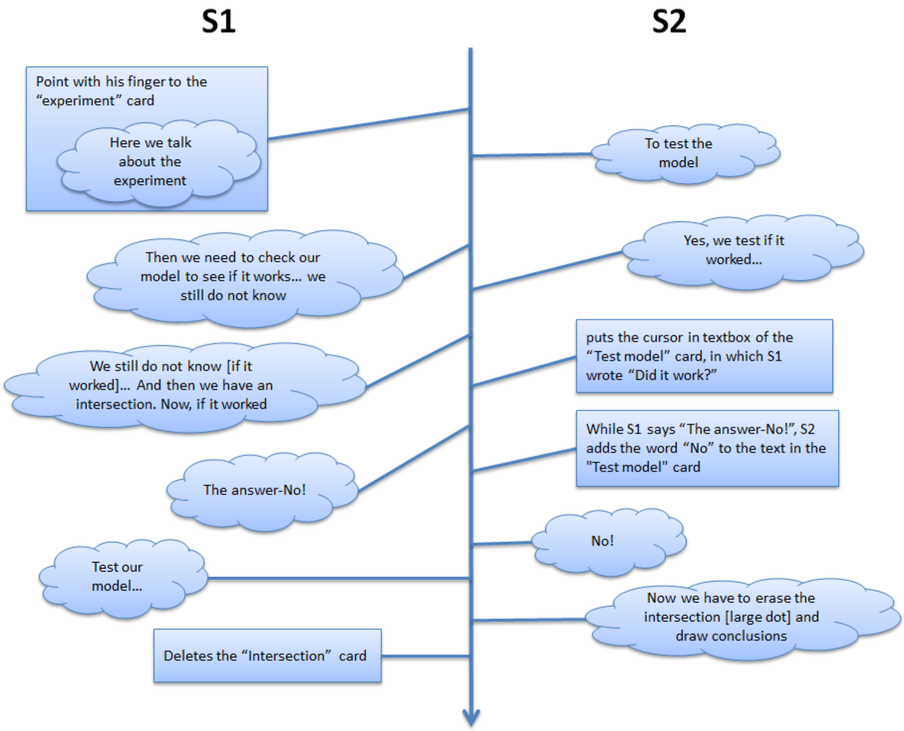


Fig. 6. Continued discussion between S1and S2 after they re-join the same computer

Discussion

We showed how the students collaboratively write down their understanding of their group work into the selected planning cards. This collaborative thinking process is externalized in their dialogues and when they work with the visual planning language. This use of the visual language affords (1) correlating their concrete challenge-based experience with abstract scientific concepts (e.g. Experiment, build a model and find hypothesis), (2) create their own concepts (*Blank card #1 - The question)* and, (3) explicate the order of these actions through their oral discussion and the plan.

Another dimension of the students' talk about scientific concepts is the expansion of their dialogic space [16] by bringing in more perspectives (*widening*) to the concept of *hypothesis* first through the students' dialogue (e.g S1 *"Find hypothesis. What is our hypothesis? "* S2 *"How about understanding the problem?",* cf. middle of Fig 5), and second as a result of this discussion putting cards #1 (define a question) and #2 (finding hypothesis) in their planning map. To this end, finding a hypothesis was given a new perspective of first defining the question and only then making a hypothesis. Expanding the dialogic space of the groups' discussions was identified as one of the characteristics of L2L2 as it implies for the students' togetherness with learning new concepts through their dialogues.

In addition, we can see how the two students interplay between past and present while they were reflecting upon their work. When trying to answer the question *What is the point that is equidistant from all vertices of any triangle*? S1 created a map in with the hypothetical situation that their model did work (although he already knew that their model–suggested solution didn't work during the previous session). In this case S1 regards their working plan to contain both actual and hypothetical actions and by that he referred correctly to the planning idea that encompass possibilities, conditions and activities meant to support the groups' road map for solving their challenge. Although he *already knows* that they found out that the *intersection of medians* is not the solution for the case of a general triangle, he still puts a decision intersection to explicitly represent a hypothetical step that the model did work as a possible step to be checked in his work plan. This approach wasn't taken by S2 as we showed earlier, because when he rejoins S1, he immediately writes in the "experiment" card #3 that the experiment that they conducted did not work. This incident pushed S1 to explain and reflect over his work to student S2 for putting both of them on the same page. We see this episode as a good example to show how the work with the Planning Tool initiated reflection over the groups' work.

5 Conclusion

In this paper we presented the Metafora approach to support students in complex multi-tool scenarios with the goal of making them able to learning to learn together (L2L2). Specifically, we concentrated on the activities of collaborative planning and reflection that guided us in the design of a dedicated Planning Tool for Metafora along the lines of the scaffolding design framework [11]. We hypothesized that the design of the Metafora learning situations together with the use of the visual language in the Planning Tool triggers a dialogue, afforded by a dialogic space, towards widening - bringing in a new perspective – as well as deepening - becoming aware of (and questioning how to explicate) a covert discourse. As shown in our classroom practice, the students' discussions around the use and visual cards widened their dialogic space and provoked meaning making processes over scientific activities and sequences.

Acknowledgments. The Metafora project is co-funded by the European Union under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme for R&D (FP7), Contract No. 257872. We thank our colleagues in the project for the fruitful discussions and cooperation to support L2L2 and especially our programmer students and our colleague Andreas Lingnau for his supportive feedback and revision of this document.

References

1. Barron, B.: When smart groups fail. The Journal of the Learning Sciences, 307–359 (2003)
2. Barab, S., Squire, K.: Design-Based Research: Putting a Stake in the Ground. The Journal of the Learning Sciences 12(3), 307–359 (2003)

3. Davis, E.A.: Prompting middle school science students for productive reflection: Generic and directed prompts. The Journal of the Learning Sciences 12(1), 91–142 (2003)
4. Ferraris, C., Martel, C., Vignollet, L.: LDL for collaborative activities. In: Luca Botturi, L., Stubbs, S.T. (eds.) Handbook of Visual Languages for Instructional Design: Theory and Practices, pp. 224–251. Information Science Reference, Hershey (2008)
5. Harrer, A., Hoppe, H.U.: Visual Modelling of Collaborative Learning Processes - Uses, Desired Properties, and Approaches. In: Luca Botturi, L., Stubbs, S.T. (eds.) Handbook of Visual Languages for Instructional Design: Theory and Practices, pp. 281–298. Information Science Reference, Hershey (2008)
6. Hoyles, C., Noss, R., Kent, P.: On the integration of digital technologies into mathematics classrooms. International Journal of Computers for Mathematical Learning 9(3), 309–326 (2004)
7. Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., Fischer, F.: Specifying Computer-supported Collaboration Scripts. International Journal of Computer- Supported Collaborative Learning 2(2-3), 211–224 (2007)
8. Loll, F.: Domain-Independent Support for Computer-based Education of Argumentation Skills, PhD Thesis, Technical University Clausthal-Zellerfeld (2012)
9. Plötzner, R., Phillipp, J., Oestermeier, U.: Organizing and externalizing activity structures in distributed learning settings. In: Hesse, F.W., Tamura, Y. (eds.) Proceedings of the Joint Workshop on Cognition and Learning through Media-Communication for Advanced e-Learning, pp. 174–179. Japanisch-Deutsches Zentrum Berlin, Berlin (2003)
10. Papert, S.: Computer-based microworlds as incubators for powerful ideas. In: Taylor, R. (ed.) The Computer in the School: Tutor, Tool, Tutee, pp. 203–210. Teacher's College Press, New York (1980)
11. Quintana, C., et al.: A Scaffolding Design Framework for Software to Support Science Inquiry. The Journal of the Learning Sciences 13(3), 337–386 (2004)
12. Suthers, D., Girardeau, L., Hundhausen, C. (2003). Deictic Roles Of External Representation in Face-To-Face And Online Collaboration. In: Proc. Of CSCL 2003 (2003)
13. Star, S.L., Griesemer, J.R.: Institutional ecology, "translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907 – 1939. Social Studies of Science 19, 387–420 (1989)
14. Stegmann, K., Weinberger, A., Fischer, F.: Facilitating argumentative knowledge construction with computer-supported collaboration scripts. International Journal of Computer-supported Collaborative Learning 2(4) (2007)
15. Botturi, L., Stubbs, S.T.: Handbook of Visual Languages for Instructional Design: Theories and Practices. Information Science Reference, Hershey (2008)
16. Wegerif, R.: Dialogic Education and Technology Expanding the Space of Learning Series: Computer-Supported Collaborative Learning, vol. 7. Springer publication house (2007)

Understanding and Supporting Reflective Learning Processes in the Workplace: The CSRL Model

Birgit R. Krogstie[1], Michael Prilla[2], and Viktoria Pammer[3]

[1] Norwegian University of Science and Technology
birgitkr@idi.ntnu.no
[2] Ruhr-Universität Bochum, Germany
Prilla@iaw.ruhr-uni-bochum.de
[3] Know-Center, Graz, Austria
vpammer@know-center.at

Abstract. Reflective learning is a mechanism to turn experience into learning. As a mechanism for self-directed learning, it has been found to be critical for success at work. This is true for individual employees, teams and whole organizations. However, most work on reflection can be found in educational contexts, and there is only little work regarding the connection of reflection on individual, group and organization levels. In this paper, we propose a model that can describe cases of reflective learning at work (CSRL). The model represents reflective learning processes as intertwined learning cycles. In contrast to other models of reflective learning, the CSRL model can describe both **individual and collaborative learning** and **learning that impacts larger parts of an organization**. It provides terminology to describe and discuss **motivations** for reflective learning, including **triggers**, **objectives** for and **objects** of reflective learning. The paper illustrates how the model helps to **analyse** and differentiate cases of reflective learning at work and to **design** tool support for such settings.

1 Introduction

Reflective learning is a mechanism to learn from experience. It plays a key role in informal learning in the workplace [1, 2] as it enables individuals and teams to handle changing work contexts and work-relevant knowledge in a self-directed manner. Reflection also helps organisations to handle competition and changing external factors [3] by enabling adaptation and fuelling innovation [4]. To achieve this, however, the outcomes of reflective learning in the organization need to be properly shared, acknowledged and utilized in relevant parts of the organization.

It has been demonstrated that reflective learning can be supported by technology [5-8] e.g. by supporting retrospective analysis, by scaffolding the reflective process, or by scaffolding the documenting and sharing of decision rationale. However, when carrying out empirical studies on reflection and its support in workplaces, we found that a solid basis to analyse reflection and conceptualize its support is missing in current research: existing models and theories for *workplace learning* do not explicitly cover reflection but have a broader focus [2], models for *reflective learning* typically

D. Hernández -Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 151–164, 2013.

focus on individual learning [9, 10], and theories describing *collaborative learning* mainly cover educational contexts [11]. Existing work demonstrating how reflective learning in the workplace can be *analysed to inform technology design* [12-15] is limited with respect to processes involving different parts of an organization.

In this paper, we contribute a model of reflective learning in workplaces, the CSRL model. It is grounded in existing theory and empirical work and supports the analysis of reflective learning at work and the design of technology to support it. The model supports unlocking the potential of reflection on multiple levels for TEL by providing a means to analyse, design and implement technology support for reflective learning.

2 Related Work: Existing Models and Approaches

2.1 Reflective Learning

We view reflective learning as the *conscious re-evaluation of experience for the purpose of guiding future behaviour* and use this term synonymously with "reflection". With this we follow Boud et al.[16], who consider reflection in the context of formal learning settings and define it as "a generic term for those intellectual and affective activities in which individuals [...] explore their experiences in order to lead to new understandings and appreciations" (Boud et al. 1985). This perspective on reflection is in line with the conception proposed by Schön [17], who in addition differentiates between reflection-in-action and reflection-on-action.

While the model presented by Boud et al. in [16] is helpful in considering how to support reflective learning, it does not explicitly link reflective learning to work [18]. To understand and support reflection in the workplace, specific characteristics of work settings need to be taken into account, most importantly the absence of teachers, the necessity for reflective learning to be highly iterative [7], its relation to other work and the various participants of reflection in practice. An important feature of the CSRL model is that it distinguishes these aspects of reflective learning.

2.2 Learning in the Workplace

Work and reflection on work are intertwined [17, 19] as reflection transforms experience from work into knowledge applicable to the challenges of daily work. Reflection has a strong social dimension and is often accomplished collaboratively by a team or working unit [20]. As a typical form of workplace learning it is often informal, while ranging from being very implicit in work to being deliberately planned with explicit goals [2]. A model for reflective learning at work needs to capture this diversity.

Models relevant for understanding and supporting learning (individual and collaborative) in the workplace can be found in CSCL, e.g., in the model of knowledge building proposed by Stahl [11]. It describes individual and collaborative learning processes as intertwined cycles and is focused on knowledge building and steps to achieve it such as clarification, negotiation and formalization. The knowledge co-evolution model by Kimmerle and Cress [21] adds an organizational perspective to this by describing how learning and knowledge building are an interplay of cognitive and social systems. However, neither of these models represents reflection and experience or include concepts relating the learning process specifically to the workplace [15].

Reflective learning has been recognized as key to bottom-up organisational learning, as it helps organizations question assumptions and to change accepted knowledge and best practice. Høyrup and Elkjær [22] argue for considering multiple perspectives on reflection in an organization, pointing out that the structuring needed from an organizational perspective must create opportunities for reflection from individualized, group and critical perspectives. Knipfer et al. [12] focus on psychological mechanisms whilst pointing out that as workplaces provide individuals with a social context, individual and collaborative learning are intertwined and must be considered together.

The CSRL model bridges the gap left by existing models of (reflective) learning, as it covers the wide range of informal learning as described above. Although it does not by itself make assumptions about organizational requirements for successful reflection, it provides a terminology and framework that can be used to describe reflection on individual, collaborative and organisational levels.

2.3 Models Supporting Technology Design for Reflective Learning

Existing models of learning processes in organizations show that insights about these processes have implications for tool use, e.g. for knowledge sharing [23]. The cyclic model of collaborative knowledge building by Stahl [11] combines individual and collaborative learning and can be a useful starting point for considering technology support for this combination, linking support to specific steps in the cycle. The model of reflective learning proposed in [14] links categories of tools to steps in a reflective learning cycle and can be used to analyse cases of computer-supported reflective learning in the workplace [13]. Continuing this line of research, the CSRL model provides a framework for linking tool use to *cycles* of reflective learning and the transitions between them, supporting reflective learning *as an iterative process*.

3 Reflection in a Hospital Setting: An Illustrative Example

To illustrate the complex nature and corresponding affordances of reflection in practice, we give an example from a real case of reflective learning at work, taken from a series of studies at different work places described e.g. in [7]. It stands proxy for a multitude of similar examples and will be used to illustrate reflection in this paper.

The example is taken from a German hospital specialized in neurological diseases, in which we observed and interviewed staff (nurses and physicians) and their practice of reflection in daily work (see [15]). We observed work on the ward for acute strokes, in which highly trained staff takes care of emergency patients. The case was selected because reflection is particularly important for medical staff in their demanding and emotionally stressful work. Also, collaborative reflection is part of the practice among colleagues, who help each other with difficult situations. Our example describes a critical incident in the work of a nurse responsible for the emergency room. The incident resulted in several reflection steps and follow-up changes of work practice.

> One day, a nurse failed to initiate the emergency procedure when a patient was brought to the emergency room and the state of the patient suddenly deteriorated severely. In cases like this, the nurse has to use the mobile telephone that he carries with her and

start an internal emergency procedure that calls a special team to the emergency room. The nurse tried to initiate the procedure, but failed and had to call the team in manually. This resulted in a time lag for the treatment, and the patient's state became critical. Although the patient recovered after this situation, the nurse felt bad about it and wanted to prevent similar situations in the future. The nurse later thought back to this situation, but did not understand what had gone wrong and why. He went to the head nurse in order to reflect with her about the procedure. The head nurse remembered similar problems in the past. Together, they found that the procedure was too complex and included too many steps to be followed in emergency situations. They came up with changes to the procedure, but were not sure whether these changes would work for others, too. The nurse and the head nurse decided to take the topic into a staff meeting to involve more people in finding a solution. In the staff meeting, some nurses reported similar problems with the existing emergency procedure. The group started to reflect on reasons for the problem and finally came up with a proposal for an adapted procedure. They decided to propose to management that the adapted procedure be used on all wards of the hospital. Also they agreed to practice emergency situations more frequently on the ward.

This example is typical for reflection we have observed, as it includes multiple cycles (e.g., individual reflection by the nurse and reflection in the group of staff), different reasons to start each of these cycles (e.g., preventing further issues) and different constraints to each cycle (e.g., in parallel to other work or in a meeting). During the analysis of this and other examples, and in designing corresponding technology, we realized that there is a need to be able to properly *describe* situations of reflection. This is needed to *identify requirements* for the support and to *design* tools and features to *implement* the support. The CSRL model was developed to suit these needs.

4 CSRL: A Model of Reflective Learning at Work

The CSRL model has been evolved from Boud's model of reflective learning [16], empirical work in different workplaces such as health, care, IT consulting, telecommunications and emergency help (e.g. [15, 24]) and a modelling process [14, 25]. The main view of the model is called the CSRL reflection cycle and will be our focus in this paper. We will refer to it as "the CSRL model" for short. In this section we establish a terminology for the key elements in the CSRL model. Then we describe the model by referring to the example in Section 3 for illustration. It should be noted that the model is understood as a framework to understand, analyse and support reflection rather than as a prescriptive, pre-defined process.

4.1 Terminology and Semantics

Reflection happens in a *reflection session*, which include a set of activities through which reflection takes place [26] and creates a reflection *outcome*. The outcome may include a change in behaviour, new perspectives and commitment or readiness for action [16]. The session is guided by a more or less explicit reflection *objective* and has reflection *participants*, which may be a single person (individual reflection) or multiple persons (collaborative reflection).

A reflection *trigger* starts the reflection as individuals or groups perceive some discrepancy, e.g. contradicting information, incongruent feelings, interpersonal conflicts and other occurrences during work, leading to a state of discomfort that the individual or group wants to overcome [12]. The reflection trigger can also be an event outside the individual that leads to awareness of the discrepancy, e.g. a reminder to review the past work week. The plan for a reflection session, including objective, participants, approach and resources, is referred to as a reflection *frame*, which may be more or less explicit and elaborate.

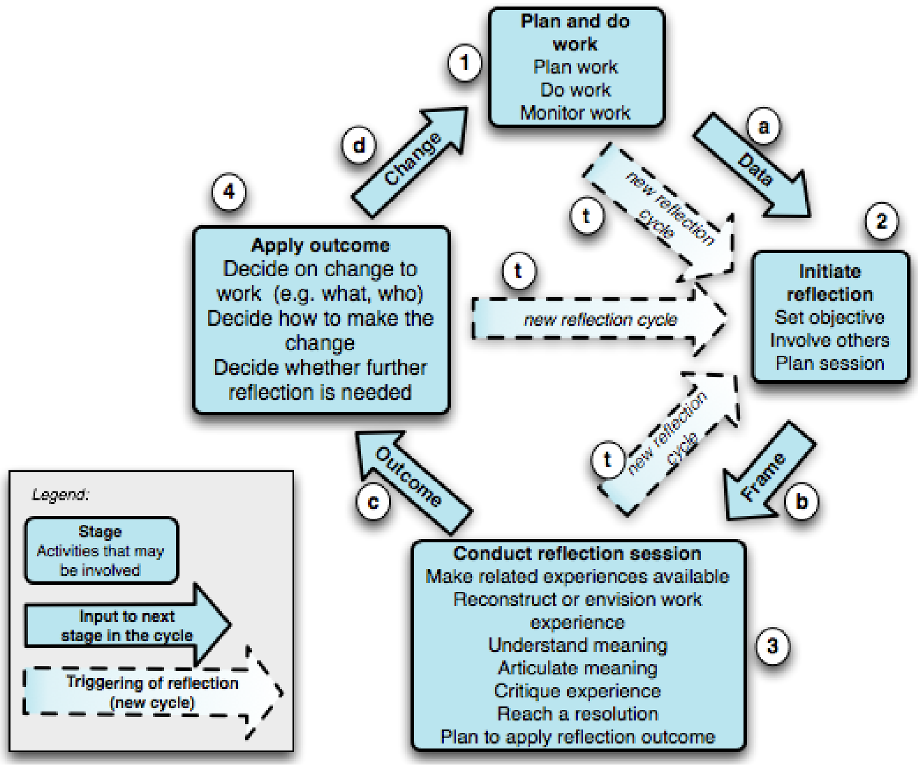


Fig. 1. The CSRL model (CSRL reflection cycle)

The CSRL model is shown in Fig. 1. It contains four stages of reflection: Plan and do work (1), initiate reflection (2), conduct reflection session (3) and apply outcome (4). The result of a stage feeds (closed arrows (a,b,c,d)) into the next stage. A reflection cycle starts with "Initiate reflection". Triggers (dashed arrows (t)) lead to this initiation. Each stage contains a non-exhaustive list of activities. In section 4.3 and 4.4, the numbers and letters (e.g. (1), (a)) refer to Fig. 1).

4.2 The Four Main Stages in the Model

The Plan and do work stage refers to conducting individual or group activity on the work arena, including everyday work, planning and monitoring. It also includes

Table 1. Stage "Plan and do work" of the CSRL model

Plan and do work (1) Input (in the context of reflective learning): *Change* **(d)** resulting from reflection Output (in the context of reflective learning): *Reflection trigger* **(t),** *Data* **(a)**	
Plan work	The particular kind of work that involves planning other work
Do work	Conduct work tasks, e.g. care for patients
Monitor work	Observe the state of the work. This includes the individual's self-monitoring during work and external monitoring e.g. by a tool

simulation of work in real or virtual environments. In the hospital example, this stage includes the episode in which the nurse fails to initiate the emergency procedure.

The stage of **initiating reflection** may be more or less elaborate. In some cases it is brief and closely integrated with work other activity. It produces a more or less explicit frame. In the hospital example, reflection is initiated when the nurse wants to understand what went wrong, when he involves the head nurse in collaborative reflection, and when the topic is brought into the staff meeting. Including the activity of involving others, this stage also provides a link between individual and collaborative learning processes [7].

Table 2. Stage "Initiate reflection" of the CSRL model

Initiate reflection (2) Input: *Reflection trigger* **(t)**, *Data* **(a)** Output: *Frame for the reflection session* **(b)**	
Set objective	Set the objective for the reflection, on the basis of the reason why reflection was triggered, e.g. understanding what went wrong
Involve others	Involve others if needed (based on the reason why reflection was triggered), e.g. asking experienced colleagues to reflect together
Plan session	Plan the session by determining the time, place and approach, e.g. conducting a meeting or approaching people spontaneously

In the **conduct reflection session** stage reflection takes place based on the frame resulting from the reflection initiation. The session has an objective, a reflection topic, it may be facilitated and involves one or more participants. The session results in an outcome. In the hospital example there are several steps characterized as reflection sessions: The nurse reflecting on why he had failed to initiate the emergency procedure, the nurses reflecting together on their experiences with this type of situation and identifying a possible improvement of the procedure or staff reflecting in the meeting.

The **Apply outcome** results in changes on work, in input to further reflection, or both. It may include involving others (e.g. team members), switching between collaborative and individual reflection or applying reflection results to work. In the hospital example, the two nurses created an improved procedure, but could not apply it

Table 3. Stage "Conduct reflection session" of the CSRL model

Conduct reflection session (3) Input: *Reflection frame* **(b)** Output: *Reflection outcome* **(c)**; *Reflection trigger* **(t)**	
Make related experiences available	Make available and share work experiences relevant to address the reflection objective, e.g. documenting a problem.
Reconstruct or envision work experience	Reconstruct relevant work or envision them as future experience, e.g. based on various data from the work process, or notes to support participants' memory
Understand meaning	Relate experience to relevant context (in light of the reflection topic), e.g. to procedures relevant for the work task
Articulate meaning	Formulate the meaning in a way that makes it understandable to others, e.g. talk about possible reasons for a problem
Critique experience	Critically evaluate the experience by use of relevant criteria, e.g. asking whether it could have been avoided and how
Reach a resolution	Agree on – or decide when there is – a satisfactory outcome of the session; formulate the outcome
Plan to apply reflection outcome	Clarify whether and how the outcome can be applied in practice

because they do not know whether other staff would consider it a good solution. The decision in the staff meeting to practice emergency situations more frequently results in changes for work practice but management needs to be included for its implementation. Outcomes may be on individual, group or organisational levels, and that both individual and collaborative reflection processes may lead to them [7].

Table 4. Stage "Apply outcome" of the CSRL model

Apply outcome (4) Input: *Reflection outcome* **(c)** Output: *Change (at the work arena)* **(d)**; *Reflection trigger* **(t)**	
Decide on change to work (e.g. what, who)	Finding out what change on the work arena should be made, and who will be involved (e.g. what scope the change should have).
Decide how to make the change	Considerations may involve whether a change to work can immediately be made and/or whether it should be described and shared, informally or formally e.g. as a recommendation
Decide whether further reflection is needed	It may be that the participants see a need for change but cannot not identify or agree on a solution, or that there is a need for the involvement of others with more expertise or power. These may be reasons to initiate a new reflection cycle.

4.3 The Transitions between the Stages in the Reflection Cycle

To support reflection as a process and not only single instances/cycles, there is a need to analyse the transitions between the stages of a cycle. The transitions involve the use of output from one step as input to the next, as shown in Fig. 1 (a,b,c,d).

Data on Work (a): Data is needed to reconstruct and make sense of work experiences. It may be more or less contextualized, more or less filtered, and more or less abstracted. The data may result from user-initiated or automatic data gathering. In the hospital example, this data could stem from notes the nurse made after the problem or from the internal call centre to reproduce which numbers the nurse had dialled.

Frame for Reflection (b): When the initiation of a reflection session is completed, the result is a frame for the reflection session. In our example, a simple frame was made up by the idea to approach the head nurse for reflection, a more elaborate frame can be seen in the meeting held with staff, which was accompanied by an agenda.

Reflection Outcome (c): A reflection session results in an outcome as described in section 4.1. The outcome may consist of several elements. It may be more or less explicit, and more or less clear with regard to implications for activity on the work arena. The outcome is the starting point for the "Apply outcome" stage. Outcomes in our example were extended internal training and a new emergency procedure.

Change on the Work Arena (d): The "change" arrow in the model depicts a change to be made or already made to work. The change has a scope in terms of what activity (tasks, processes) is affected and in terms of who (individuals, roles) are affected. A change done in the hospital example was that the nurses started to more frequently practice emergencies.

4.4 Initiating Reflection Sessions: Transitions between Cycles

The example describing reflection on the failed emergency procedure described in section 3 contains three interconnected, sequential reflection cycles (see **Fig. 2**):

Cycle 1: Individual Reflection of the Nurse: After the critical incident, the nurse thought about reasons for it to prevent similar situations in the future, but could not come up with a better understanding of the situation.

Cycle 2: Reflection Between Nurse and Head Nurse: As a result, the nurse asked the head nurse to reflect on the issue together. They identified that the emergency procedure was too complicated, but were not sure about this.

Cycle 3: Reflection in Staff Meeting: Again as a result, the head nurse decided to reflect the issue with the whole team in order to make them aware of the complexity in the emergency procedure and get their feedback.

In addition, but not observed in the study, the nurses later initiated a **fourth reflection cycle**, in which they included the hospital management, to reflect on changes on the hospital emergency procedure. **Fig. 2** shows these interconnected reflection cycles and depicts the importance of the "Initiate Reflection" stage of the CSRL model: Although it is not always an explicit and scheduled process – for example, the nurse decided spontaneously to ask the head nurse for support – people need to find proper partners and to transfer the data and information to them when they initiate a new cycle. This was easy for the nurse in the example, as she knew the head nurse was very experienced and had already heard of the problem with the emergency procedure. In other cases, the initiation of reflection with adequate resources is more challenging, as in the example when the topic needs to be communicated to all staff.

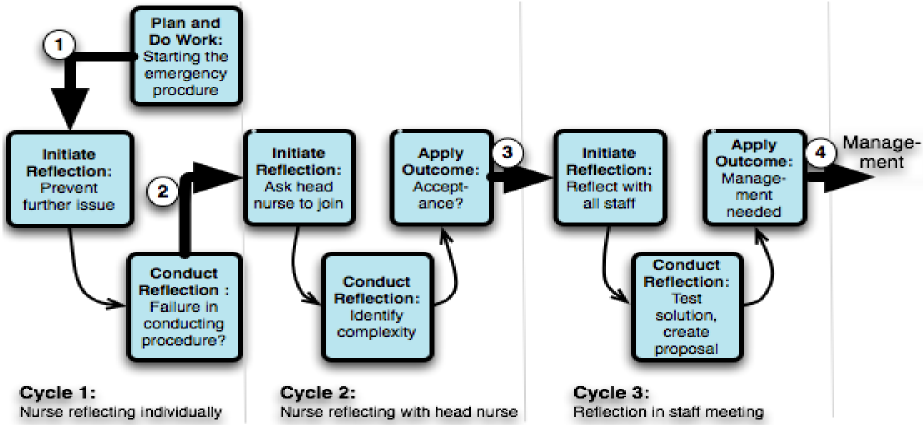


Fig. 2. Interconnected reflection cycles based on the example from the hospital

The expressiveness and added value of the CSRL model for the analysis and design of reflection at work becomes evident in **Fig. 2** as it includes three explicit transitions between cycles and thereby also between levels of reflection, that is, whether an individual reflects on her own or as part of different groups [24]. *Initiation from work* is the most intuitive start of a reflection cycle (see (1) in **Fig. 2**). This is the type of initiation of reflection most prominent in existing literature. The CSRL model adds the possibility to start a new cycle *during reflection* (2) and *after applying outcomes* of a reflection session (3). For the former, this indicates that reflecting actors need to **recognize when their reflection does not lead to an applicable outcome** and start a new reflection cycle. For the latter, it shows that if an **outcome cannot be applied** e.g. due to lacking expertise or power to do so, an additional reflection cycle, including participants with the means to apply the outcome, can lead to change as a result of reflection. (4) indicates the start of reflection by management, outside our story.

As described above, triggers initiate reflection cycles. Triggering can happen during work and during reflection, resulting in the initiation of a new reflection cycle. In order to understand how people reflect in practice and how to create support for this, we need to understand why they initiate reflection cycles and switch between them, i.e. why they start new reflection cycles as described above. For this purpose, the CSRL model includes triggers as described above (section 4.1) for the transitions to new reflection cycles. In an earlier analysis of studies conducted in three different workplaces, we identified three categories of work-related reasons leading to transitions between cycles, which are covered in the CSRL model [27]:

1) Seeking clarification or resolution: Most often, a new reflection cycle was initiated when an individual or a group needs (additional) input and support by others in order to clarify an issue or to come up with ideas for solution. In our study, this either occurred *during work*, e.g. when individuals performed badly in a certain task such as in the example with the failed alarm procedure, or *during reflection* sessions, when the reflecting actors realized that they needed additional participants (e.g. all staff) to understand the problem better and create a solution.

2) Seeking support for solution implementation: In multiple cases, we found that reflection partners were added and a new reflection cycle was started *when the reflecting actors created an outcome*, but could not implement it by themselves due to lacking expertise, power or some other issue. In the example, such a cycle was initiated after the nurses had come up with the idea to change the hospital's emergency procedure and wanted to reflect together with the responsible manager on this.

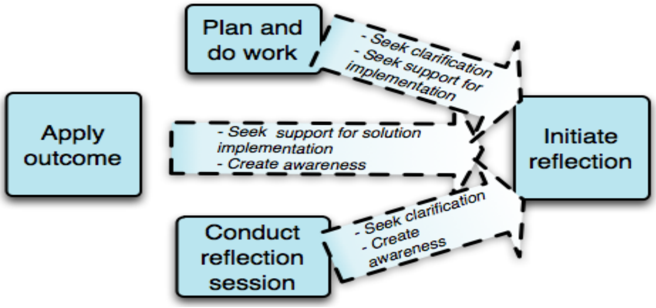


Fig. 3. Triggering of new reflection cycles in the CSRL model

3) Creating awareness: As a third category, we found reflection cycles to be initiated in order to make the reflection partners aware of a problem and to support them in learning about it. Reflection was then used to contextualize a problem by asking others to share experiences on it and to discuss options from earlier reflection cycles. This kind of triggering of reflection mainly occurred *during reflection sessions* and *after outcomes of reflection had been created*. In the hospital example, it happened when the head nurse started reflection in a meeting with all the ward staff to make them aware of the problem with the emergency procedure and to test drive the solution she and the other nurse had come up with.

Fig. 3 shows these triggers in the context of the CSRL and illustrate that they explain how and why new reflection cycles are initiated.

5 Supporting Design and Implementation of Technology for Reflection with the CSRL Model

Besides the descriptive and analytical strengths of the CSRL model, its main contribution to TEL is that it informs the design and implementation of technology supporting reflective learning at work. In an initial step of developing the model, its main stages and their activities were **mapped against technology** available to support individuals and groups in gathering data, exchanging it, communicating about and sustaining results [14, 25]. This enables designers to use the model for analyzing and describing cases of reflection – by analyzing reflection cycles as demonstrated in this paper – and **choose proper technology** support for activities important in the respective case.

This has already been applied to cases in which for example the tracking of people's mood was central for the reflection of meeting facilitation [28], in which the documentation of critical situations and the articulation of different perspectives in these situations was needed for reflection [24] or in which the reconstruction of work experience was crucial for reflecting on it [29]. Addressing the explicit linking of the CSRL reflection cycle to the use of reflection tools can be considered a distinct view of the CSRL model and will be subject to an upcoming publication.

Besides this direct mapping between activities of reflection and meaningful technology support, the connected cycles of reflection and reasons for starting cycles that we described with the CSRL model create additional information and requirements for the design of reflection tools. First of all, they show that reflection tools not only need to provide data on work to start and feed reflection, but also need to **create, exchange and maintain a context of reflection** throughout different cycles. This context includes outcomes of earlier cycles, which can be used as input or topics for later cycles, communication on the reflection topic that is relevant for the cycle, as well as experiences and understanding shared in previous cycles to support sense making in the next cycles. The context needs to be shared across cycles to ensure that reflection does not have to be started from scratch or with high efforts to recreate the context (e.g., if there is little time between two cycles or if there are outcomes from earlier cycles).

Looking at the **transitions and triggers** connecting steps and cycles in the CSRL model, it becomes clear that reflection tools also need to provide support to keep the cycle active until a resolution has been reached and applied. For the transitions described in the model (data, frame, outcome, change; see section 4.4) this means reflection tools need to create an output for each phase that feeds into the next.

Data: Data such as stress level figures, work performance measures or notes on experiences form the basis of reflection by supporting human memory [30]. Data needs to be available to individuals and groups: Writing down and sharing experiences as in the example of the nurse enables more objective reflection on past episodes.

Frame: Tools need to support people in creating a frame for reflection, which needs to be transferred to the reflection session. This frame may differ: In our example, the nurses created an agenda for a meeting as a frame, but they could as well have written down the experience and broadcasted a questions such as "Can someone think of a solution for this" in a tool instead of starting a meeting.

Outcome: For reflection to be successful and take multiple cycles, the preservation and exchange of outcomes is critical – without tool support, outcomes are easily forgotten or not shared with others [7]. Reflection tools should enable the documentation of outcomes, including a link to the data reflected on to contextualize the outcome (see also [24]) In the case of the nurses, the decisions to change the amount of training done and to try to change the overall emergency procedure could have been documented in a tool to follow up on them later or ask others whether they agree.

Change: Change needs to be applied on the level of work (processes). Reflection tools could support this by creating proposals for change that can be tracked. Data stemming from this tracking can then be input for the reflection of the changed work procedures. In our example, a tool could have created the proposal to change the procedure, which could be checked regularly for its implementation and effect on work.

Concerning the triggers discussed in section 4 the model emphasizes the need for **supporting the phase of initiating reflection**: Although this phase is often implicit when there is no dedicated planning session for a reflection cycle, there is a need for support in involving reflection partners. Moreover, the initiation of new reflection cycles from all steps of the CSRL model suggests that reflection tools should support awareness on or assessment of the state of reflection in different steps, e.g., whether an outcome was created that was agreed upon by a group or whether a solution can be applied. If this assessment has a negative result, the tool should support users to identify other reflection participants accordingly.

6 Conclusion

In this paper, we have stressed the descriptive power of the CSRL model to represent aspects of reflection in the workplace that are currently not well covered by existing frameworks, and that need to be understood in order to provide appropriate technology support. The model advances the state of the art as **it anchors reflection in work practice** and describes it as part of such practice, in that it includes **individual and collaborative reflection cycles**, in that it explains **transitions between cycles** and **reasons** for these transitions, and in that it informs technology support for reflection beyond individual cycles. This gives the model various potentials in TEL:

With the CSRL model, reflective learning in the workplace can be represented as a process involving multiple, interconnected cycles. This is crucial to understanding reflective learning because the final reflection outcome and its application as change on the work arena may depend on the opportunity to initiate additional cycles. The descriptive power to model multiple cycles allows a detailed analysis of a case which can help analysts/designers understand how multiple cycles are, or could be, involved. This can help them create a more complete model of the case.

With the CSRL model, each reflection cycle can be characterized in a way that captures important differences between cycles. This can be done by describing the elements in the cycle (stages, transitions) and by identifying patterns in types of cycles or interconnected cycles (e.g. individual sense making followed by consulting a peer, or reflection in a team followed by involving management).

With the CSRL model, tool support can be adapted to the particular characteristics of the reflective learning process in question. Examples of considerations include how to support the involvement of other parts of the organization, how to bring about actual change with a particular organizational scope (e.g. one's team, the entire organization), how to help a team come up with creative solutions in a reflection session, how to help a learner become aware of the opportunity to reflect.

To fully exploit the potential of the model, further work is needed, especially in understanding and describing **single reflection cycles and their relationship**. Research building on the contribution presented here should apply the CSRL model to different types of cases of reflective learning, to further explore its descriptive power across different domains. More detail should be provided on the dimensions along which model element can be characterized. Also the connection between reflection cycle elements and tool support should be detailed further, as a step towards making the CSRL model a design framework that connects the characteristics of reflective learning processes to recommended tools and tool features.

Acknowledgements. The MIRROR project is funded under the FP7 of the European Commission (project number 257617). We would like to thank our colleagues for their contributions and valuable input during discussions.

References

1. Boud, D., Garrick, J.: Understanding learning at work. Routledge (1999)
2. Eraut, M.: Informal learning in the workplace. Studies in Continuing Education 26(2) (2004)
3. Cressey, P.: Collective reflection and learning. In: Boud, D., Cressey, P., Docherty, P. (eds.) Productive Reflection at Work. Routledge (2006)
4. Høyrup, S.: Employee-driven innovation and workplace learning: Basic concepts, approaches and themes. European Review of Labour and Research 16(2), 12 (2010)
5. Krogstie, B.R., Divitini, M.: Supporting Reflection in Software Development with Everyday Working Tools. COOP. Springer, Aix-en-Provence (2010)
6. Lin, X., et al.: Designing Technology to Support Reflection. Educational Technology, Research and Development 47(3), 43–62 (1999)
7. Prilla, M., Pammer, V., Balzert, S.: The Push and Pull of Reflection in Workplace Learning: Designing to Support Transitions between Individual, Collaborative and Organisational Learning. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 278–291. Springer, Heidelberg (2012)
8. Xiao, L., et al.: Promoting Reflective Thinking in Collaborative Learning Activities. In: Eighth IEEE International Conference on Advanced Learning Technologies (ICALT). IEEE, Santander (2008)
9. Boud, D., Keogh, R., Walker, D.: Promoting Reflection in Learning: A Model. In: Boud, D., Keogh, R., Walker, D. (eds.) Reflection: Turning Experience into Learning, pp. 18–40. RoutledgeFalmer (1985)
10. Kolb, D.A., Fry, R.: Towards an applied theory of experiential learning. In: Cooper, C.L. (ed.) Theories of Group Processes, pp. 33–58. John Wiley, London (1975)
11. Stahl, G.: Building collaborative knowing. In: Strijbos, J.-W., Kirschner, P.A., Martens, R.L. (eds.) What We Know About CSCL And Implementing It In Higher Education. What we know about CSCL, pp. 53–85. Kluwer Academic Publishers, Boston (2002)
12. Knipfer, K., et al.: Reflection as catalyst for organisational learning. Studies in Continuing Education 35, 19 (2012)
13. Krogstie, B., Prilla, M.: Tool support for reflection in the workplace in the context of reflective learning cycles. CEUR Workshop Proceedings, vol. 931, pp. 57–71 (2012)
14. Krogstie, B., et al.: Computer support for reflective learning in the workplace: A model. In: International Conference on Advanced Learning Technologies (ICALT) 2012. ACM, Rome (2012)
15. Prilla, M., et al.: Individual and Collaborative Reflection at Work: Support for Workplace Learning in Healthcare. In: CSCL 2013. ISLS, Wisconsin (2013)
16. Boud, D., Keogh, R., Walker, D.: Reflection: Turning Experience into Learning. RoutledgeFalmer (1985)

17. Schön, D.: The Reflective Practitioner. Basic Books, Inc. (1983)
18. Boud, D., Cressey, P., Docherty, P.: Productive Reflection ad Work: Learning for Changing Organizations. Routledge (2006)
19. Lave, J.: The practice of learning. In: Chaiklin, S., Lave, J. (eds.) Understanding Practice: Perspectives on Activity and Context, p. 20. Cambridge University Press, Cambridge (1993)
20. Høyrup, S.: Reflection as a core process in organisational learning. Journal of Workplace Learning 16(8), 13 (2004)
21. Cress, U., Kimmerle, J.: A systemic and cognitive view on collaborative knowledge building in wikis. Computer-Supported Collaborative Learning 3, 105–122 (2008)
22. Høyrup, S., Elkjær, B.: Reflection - beyond the individual. In: Boud, D., Cressey, P., Docherty, P. (eds.) Productive Reflection at Work. Routledge (2006)
23. Kimmerle, J., Cress, U., Held, C.: The interplay between individual and collective knowledge: Technologies for organisational learning and knowledge building. Knowledge Management Research & Practice 8(1), 12 (2010)
24. Prilla, M., Degeling, M., Herrmann, T.: Collaborative Reflection at Work: Supporting Informal Learning at a Healthcare Workplace. In: GROUP 2012. ACM, Sanibel Island (2012)
25. Krogstie, B.R., et al.: Collaborative Modelling of Reflection to Inform the Development and Evaluation of Work-Based Learning Technologies. In: i-KNOW 2012. ACM ICPS, Graz (2012)
26. Pammer, V., et al.: Reflective Learning at Work - A Position and Discussion Paper. In: ARNets11- Awareness and Reflection in Learning Networks, Palermo, Italy (2011)
27. Prilla, M., Pammer, V., Krogstie, B.: Fostering Collaborative Redesign of Work Practice: Challenges for Tools Supporting Reflection at Work. In: Proceedings of the European Conference on Computer Supported Coopertive Work, ECSCW 2013 (in press, 2013)
28. Fessl, A., Rivera-Pelayo, V., Pammer, V., Braun, S.: Mood Tracking in Virtual Meetings. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 377–382. Springer, Heidelberg (2012)
29. Mora, S., Boron, A., Divitini, M.: CroMAR: Mobile Augmented Reality for Supporting Reflection on Crowd Management. International Journal of Mobile Human Computer Interaction 4(2), 14 (2012)
30. Knipfer, K., et al.: Computer Support for Collaborative Reflection on Captured Teamwork Data. In: 9th International Conference on Computer Supported Collaborative Learning (2011)

Rhythm Reading Exercises with PWGL

Mika Kuuskankare[1] and Vesa Norilo[2]

[1] DocMus–Department for Doctoral Studies in Music
Sibelius Academy, Finland
mkuuskan@siba.fi
[2] Centre for Music and Technology
Sibelius Academy, Finland
vnorilo@siba.fi
http://www.siba.fi/PWGL

Abstract. This paper presents rhythm reading, one of the elementary ear training exercises, as a pedagogical software application of PWGL. We use different kinds of stochastic and mathematical models to generate a rhythmic database. The database is divided into several categories, including, binary or ternary, euclidian, afro-cuban, corpus-based, and contemporary. Our musical constraints systems is used to define a rule set, which, in turn, can be used to automatically generate graded rhythm reading exercises. The user is then presented with a musical score, and he or she can perform a reading with any percussive instrument or voice and a microphone connected to a computer. Our novel signal processing system is utilized to analyze the reading. Finally, visual feedback and statistics are displayed directly as a part of the exercise. In this paper we present our rhythm reading application, and discuss the details of its implementation.

Keywords: Educational technology, music education, interactivity, multimedia learning environments, human-computer interaction, constraints, signal processing.

1 Introduction

Today, computer-based music tuition is a growing field of interest. Several studies have found that Computer Assisted Instruction (CAI) is in general beneficial for students and that the learning gains are substantial (see for example Hofstetter, [2]). Furthermore, [3] concluded that the use of CAI in teaching rhythm dictation is especially effective.

There are numerous web applications that are targeted at basic-level music education, including those at `www.emusictheory.com/` and `www.teoria.com/`. One advantage of web applications is that they work on most browsers and on a majority of the operating systems. There exists also traditional computer software aimed specifically at Computer Assisted Music Instruction (CAMI). However, most of these applications are not extendable, they do not provide the tools for users to prepare their own educational material and they tend to focus on one

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 165–177, 2013.

specific problem such as instrument tuition, as is the case with the IMUTUS[11] system. One of the first CAMI systems was the University of Delaware's Graded Units for Interactive Dictation Operations (GUIDO). The GUIDO system taught the identification of intervals, melodies, chord qualities, harmonies, and rhythms. One of the more recent projects is LenMus (`http://www.lenmus.org/`). which is an open-source software for developing applications for practicing music theory and aural skills. There are also other platforms, for example OpenMusic, which are not specifically geared towards pedagogy but which nevertheless can be used for that. Furthermore, the use of the combination of Strasheela and GNU/Lilypond for generating rhythmic exercises has been reported by Percival et al.[10] However, they concentrate on developing a rule set for generating the exercises, rather than developing an application that makes it possible to input the solution, check it automatically, and give feedback to the user, as is the case in this paper.

PWGL[8], is a visual computer language designed for the applications of computer-assisted composition, music theory and analysis, software synthesis, and music notation. PWGL has mostly been used as an artistic/compositional tool. In our case, all of the techniques that we have developed lend themselves well to the production of teaching material, and, also, provide a basis for the automatic analysis of user input. PWGL should prove to be an ideal platform for realizing music tuition applications as it incorporates a flexible real-time music notation engine, a rule-based scripting language, a multi-rate capable sound synthesis engine, and a content authoring tool (CAT). Our music notation package, ENP[5], can be used both as an input and output tool. Most of the web-based applications rely on textual input. ENP, however, can be used as a GUI component for representing musical data. The Kronos[9] synthesis engine can be used in signal processing tasks, such as pitch estimation, or, as in our case, for transient detection. Synthesized output, in turn, can be used in advanced ear-training drills, where, for example, microtonal pitch resolution is required. Our rule-based composition/scripting system[6] can be used to generate drills, and, also, to check the user input and mark the mistakes directly in the score. Finally, our content authoring tools make it possible to arrange visual and textual components to create interactive training content. The material produced using our system can be further converted into other formats, such as, PDF or HTML, which makes it possible to work and study offline as well. In the case of PDF, for example, the drill sections would obviously not be interactive, but those sections would nevertheless be exported and they would contain material, such as blank staff lines for sketching the potential solution, that make sense in paper form. Therefore, the non-interactive version can also be used as study material.

This paper marks the start of a series of studies where we focus on implementing pedagogical material with the help of PWGL. Our long-term plan is to extend the repertoire to cover several aspects of music tuition ranging from ear training to counterpoint. One of the benefits of PWGL is that it is a fully programmable and extendable system. Through its visual interface it is possible for the teachers, even with no considerable programming background, to develop independently drills and study material for music theory, music analysis and composition, etc. Many studies concentrate on the theoretical aspects of learning but only a few look at the practical issues involved in developing CAMI software.

In the remainder of this paper we present one specific pedagogical application, the rhythm reading, and discuss its implementation in detail. The generative part, realized with the help of musical constraints, takes care of generating suitably graded rhythmic exercises and is covered in Section 2. The synthesis part, which assumes responsibility for detecting the user input and informing our host application, is covered in Section 3. The analysis of the user input is discussed in Section 4, and, finally, the application part, which presents users with a graphical user interface and provides for the visual feedback, is discussed in Section 5.

2 Generative Part

In our pedagogical application we use several techniques to produce teaching material. The general strategy is that we first use some sort of stochastic or mathematical system to create a database of rhythmic figures. Our constraints system, is, then, used to implement a set of rules that guarantees coherence, a sense of development, the feasibility of the musical phrases, and the ability to control the level of difficulty. Our database is divided into several categories: (1) binary, (2) ternary, (3) euclidian, (4) afro-cuban, (5) corpus-based, and (6) contemporary.

The idea behind this type of classification is that the groups are arranged in an increasing order of complexity/difficulty. Binary and ternary rhythms present entry level material where the number of possible rhythms is limited and the rhythms are simple. The rhythms in the euclidian, and afro-cuban categories, are all generated using the general idea of euclidian rhythms[13]. The corpus-based category incorporates a set of constraints that are automatically generated by computationally analyzing a musical corpus. Finally, the contemporary category includes both irregular rhythms and meters, as well as almost unrestricted combinations of them. These are generated, for example, using the 'mode de valeurs' system by Messiaen and other systems coming from mathematics, such as the Golomb ruler, which results in maximally irregular rhythms.

2.1 Generating the Exercises

Due to space constraints, we choose, as an example, the binary rhythm group. It is not our aim to develop a comprehensive set of rules, but, instead, to give a working example of the implementation.

A database of acceptable rhythms is first generated. The method is relatively simple and is not covered here. The database is represented using the PWGL RTM-notation. In its entirety it is as follows: (-1) (-3 1) (-1 1) (-2 1 1) (-1 3) (-1 2 1) (-1 1 2) (-1 1 1 1) (1) (3 1) (1 1) (2 1 1) (1 3) (1 2 1) (1 1 2) (1 1 1 1).

The notation is quite straightforward: Positive numbers denote the length of a note in units. Negative numbers, in turn, represent rests. The base duration chosen here is one quarter-note, i.e., the list of numbers represents the internal distribution of durations inside a quarter-note. Thus, the encoded rhythm (3 1) in a music notation is 3/4 of the base unit, i.e., a dotted 8th-note, followed by 1/4 of the base unit, which, in turn, is a 16th-note. An example using common music notation is given in Figure 1. Note that the sum of all durations in the list is 4, thus the ratios 3/4 and 1/4.

(3 1) = 

Fig. 1. The encoded rhythmic figure (3 1) and the resulting rhythm in common music notation. The sum of all the elements in the list is 4, and the base duration chosen here is a quarter-note. Therefore, the list (3 1) results in the ratios 3/4 and 1/4, which, in turn, result in a dotted 8th-note (three quarters of a quarter-note) and an 8th-note.

Next, we examine in detail the rules used to produce the exercise, using the database of rhythmic figures defined earlier. The rule set consists of three relatively simple rules. The first rule (line 1) controls the first rhythm in the exercise. Here, we make sure that the first event of the exercise is a note instead of a rest. This is mainly because of our input scheme: the reading is started as soon as the system detects the first transient, as explained in section 4.1.

The next rule (line 3) states that we want to avoid repetition in the space of three consecutive rhythms. And, finally, lines 5–10 control the structure of the whole exercise: The last beat is constrained to (1), that is, a quarter-note. Also, every fourth figure must begin with a longer duration. This is to create a cyclic feel.

```
1   (i1 (?if (plusp (first i1))))

3   (* ?1 ?2 ?3 (?if (setp (list ?1 ?2 ?3) :test #'equal)))

5   (* ?1 (?if (if (= (cur-index) (cur-slen))
6                  (equal ?1 '(1))
7                (if (zerop (mod (cur-index) 4))
8                    (or (eql ?1 '(1))
9                        (> (car ?1) 1))
10                 (neql ?1 '(1))))))
```

The rule syntax is covered in detail, for example in [7], and cannot be discussed in depth here. However, Figure 2 illustrates the relationship between the rhythmic figures in the database and the different components of the rules (the rule is the same as in Listing 2.1, line 3). Here, the rule is divided between two lines. The topmost line shows the pattern-matching part and the bottommost line, in turn, shows the rule part. The current train of rhythms, i.e., the current partial solution, is shown in the middle. As can be seen, we can make a symbolic link between the composed rhythmic score and the rules (shown by the connections a–c). This information is then used by the rule to apply certain constraints as demonstrated by the connection (d) in Figure 2. Here, we maintain that there can rhythmic figure cannot be repeated inside a window of three consecutive figures. Note also, that in this case the rule would not accept this solution as the first rhythm (a) is repeated in (c). Again, these three rules are given here as an example. In reality, there would be a collection of rules for generating exercises of different levels of difficulty.

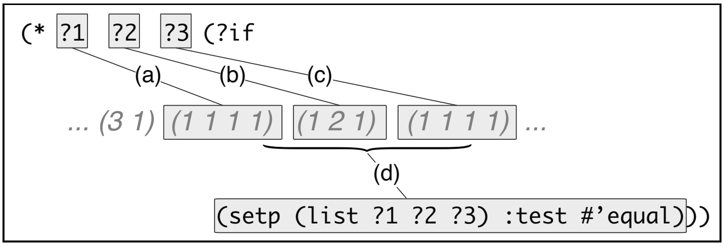


Fig. 2. A snapshot of the process of applying constraints. Here, the current solution is displayed in the middle. The rule is wrapped around it to make it easier to visualize the connections between the solution and the parts of the rule. Note, that in this case the rule would fail because the first and the third rhythms are the same (see a and c). This is the scenario that is explicitly forbidden by this rule (d).

3 Implementation of Machine Listening

Machine listening is essentially signal processing. Such tasks are not suited for implementation in PWGL, mainly because it is not highly efficient in numerical processing. Traditionally, a special component called PWGLSynth has been used as a domain-specific tool for signal processing tasks. At the moment, it is in the process of being replaced by *Kronos*[9], an emerging signal processing language.

Kronos aims to be a general, yet easy-to-use platform for realizing customized signal processors. It is particularily relevant to the task at hand by is its comparatively superior capability for multi-rate processing, especially in the case of audio analysis.

The backend implementation of *Kronos* is a just-in-time compiler[1]. In this application, the signal processor is described textually and compiled into native machine code on the fly to enable high numerical performance and real time-signal processing. The compiler runs as a standalone *Kronos* server process, communicating with client processes via Open Sound Control (OSC)[14], an open standard for musical signals in a networked environment.

OSC is used to transfer the signal processor algorithm from PWGL to *Kronos* server, which subsequently compiles and executes it, transmitting the analysis data back to PWGL via another OSC channel. Because OSC is network transparent, the communication between PWGL and *Kronos* can take place equally well whether the two programs run on a single computer or separate computers on a network.

Since all the communication is based on an open standard, third party software can easily extend or automate the process by intervening in the OSC communication channel.

3.1 Signal Processing Implementation

In this section, the *Kronos* program and the algorithm used to analyze the musical performance is briefly explored.

To detect musical onsets, the raw audio signal from a microphone is gradually refined towards a higher level of abstraction. Initially, the envelope or the momentary loudness of the signal is extracted from the raw vibration. Secondly, transients, or rapid increases in loudness are detected to find potential musical onsets. Finally, onsets occurring in too rapid a succession are inhibited to prevent multiple detections of a single actual onset. While the algorithm is hardly novel, it is covered here to demonstrate the relative ease of the *Kronos* implementation.

Envelope Estimation. A well-known method for envelope estimation is the use of root mean square technique. The momentary loudness is obtained from a finite duration or a window of time around the point. By extending the window, the loudness estimation is more robust against the fluctuation caused by low frequency vibration. However, the envelope estimation exhibits increasing lag in responding to transients. This property turns out to actually benefit the subsequent transient analysis step.

Moving into *Kronos* code, a traditional method to optimize RMS calculation is employed. The algorithm involves integrating over a rectangular window of the squared audio signal. Invoking mathematical commutativity, the rectangular window can be substituted with its derivative - positive and negative pulses corresponding to the left and right edges of the window - followed by double integration. This optimization allows the window to be computed with a simple subtraction, decoupling the computational cost from window size. An optimized RMS integrator stage is presented as a *Kronos* function shown in Listing 1.1. The window functions are shown in Figure 3.

The output of this function represents the integral of the squared audio signal over a window of length *window-size*. By adjusting this parameter, the lag and stability of the estimator can be controlled. What remains is to divide the integral by the window size and to take the square root.

Transient Detection. The estimation lag of the RMS method can actually be leveraged to build a transient detector. Having two envelope estimators with different lag characteristics, a rapid increase in loudness can be detected by contrasting the faster estimator with the slower one. The behavior of the estimators is illustrated in Figure 4.

When the estimators encounter a transient, the fast estimator rapidly reaches the new loudness while the slow estimator takes longer to reach it. Whenever the fast estimator exceeds the slow estimator, a transient is underway. The sensitivity

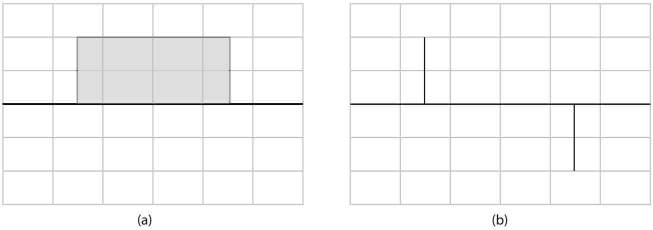


Fig. 3. A Rectangular window (a) and its derivative (b)

Listing 1.1. RMS Integrator implemented in *Kronos*

```
Integrator(sig window-size){
        squared = sig * sig
        energy = z-1('0 energy) + squared - Delay(squared
           window-size)
        Integrator = energy
}
```

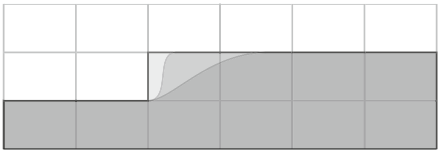


Fig. 4. Actual loudness, fast and slow estimator behavior

of the detector can be adjusted by specifying a ratio by which the fast estimator value must exceed that of the slow estimator to trigger a transient.

The transient detector function, shown in Listing 1.2, makes use of two *Integrators* of different window size from Listing 1.1. An optional decimation is used to decrease the signal rate at this stage to spare computational resources consumed by the division and square root operations.

Onset Analysis. CD-quality digital audio consists of 44100 numerical samples per second. The goal of the signal analysis step is to dramatically reduce the amount of information, ending up with just a single event per detected musical onset.

Listing 1.2. Transient detection utilizing two different RMS Integrators

```
Transient(sig bigwindow smallwindow decimation)
{
    (slow fast) = Reactive:Decimate(
                            (Integrator(sig bigwindow)
                             Integrator(sig smallwindow))
                             decimation)

    Transient = Sqrt((fast + 0.0001) / (slow + 0.01))
}
```

Listing 1.3. Onset analysis, initial gating to drop non-transients

```
transient = RMS:Transient(Audio:Mono-Input(#1) #2000 #64
    #16)

peak = Reactive:Gate(transient transient > 0.65)
time = Coerce(Float Reactive:Time())
detected = (time peak)
```

The first information reduction step is achieved by utilizing *Kronos*'s signal decimation facility, built into the transient detection function shown in Listing 1.2. In the example implementation, a factor of 16 is used to bring the signal rate of the envelope down from 44.1kHz to 2756Hz. This reduces the computational cost of subsequent analysis steps, while retaining a very detailed picture of the loudness behavior of the signal.

In the next step, all samples of the transient curve below a certain threshold are dropped by a Gate operator. Any samples passing this gate are associated with a time stamp describing when they occurred.

Finally, the detected transient is contrasted to the previous detection. If less than 10 milliseconds have passed, the transient is rejected as likely belonging to a single musical event.

Listing 1.4. Onset analysis, final gating to remove spurious duplicates

```
previous-detected = z-1('(0 0) detected)
(prev-time prev-level) = previous-detected

Onset-Analyzer = Reactive:Gate(detected time > (prev-time
    + 0.01))
```

4 Analyzing the User Input

Once the user has performed the reading the user input is analyzed and the analysis is displayed directly in the score.

4.1 The Metrics

During a signal processing run, the *Kronos* component of the performance analyzer has transcribed a set of onset events, which in their numerical form consist of time-loudness pairs.

To facilitate analysis, the detected sound events must be associated by the generated score, such as those described in Section 2. This is a fuzzy process, as there is no certainty that the performer has hit all the notes nor that the detection algorithm works correctly in all possible cases. Further, there is no absolute tempo reference, so the program needs to deduce which sound events correspond to which score events.

An estimated match can be computed by choosing an anchor event; this sound event is assumed to have correct timing and duration. A tempo can be set by choosing a score event that the anchor event represents – this becomes a relationship between absolute and musical duration. The quality of the match can be assessed by computing prospective absolute timing for all the events in the score. Each "ideal" event is contrasted with the nearest sound event detected from the reading, and the relative timing error is stored. The cumulative timing error is the total *badness* of the match.

All possible matches are selected by considering each detected sound event as each score event. The best matches are, in all practical cases, those where the anchor event was linked to a correct score event. An averaged tempo is obtained from several of the best matches, and final timing errors can be computed by utilizing the anchor/score link of the best match along with the averaged tempo.

This scheme works well with relatively short exercises. It is notable that the complexity of the algorithm is $O(N^2)$ in relation to the number of events in the exercise. As a future expansion, the matching should be done in smaller sections. This has two benefits: firstly, the number of computed matches remains the same with longer exercises and secondly, the tempo estimation prioritizes time-local musical context rather than averaging over the entire exercise.

4.2 Visualizing the Metrics

We use several different ways to visualize the data obtained via our metrics (see Figure 6 for examples).

The relative error between the correct rhythm and the user input is displayed below the rhythmic figures. The small colored note head indicates the distance from the actual event. Also, a graph is drawn below the note heads. This provides a continuous view of the user input. The graph is drawn so that a positive error is drawn above and negative error is drawn below the 0-axis line. The height of the graph indicates the size of the error. Furthermore, a color scheme is used with interpolating colors: green indicates either no error or a minor error and red indicates a major inference. The resolution and the granularity of the scale can be adjusted by the user, i.e., the teacher, according to the level of the student or students in question.

All of the information described above is provided with a 'temporal metrics expression'– an ENP expression specifically developed for this purpose. ENP expressions are Lisp-based multipurpose dynamic graphical objects that can be used to represent different kinds of information as part of a musical texture. How to implement expressions in ENP is described in [4] and is not covered here.

Visualizing the results of our metrics in the score is a relatively simple procedure. For each event we know the displacement of the user input in seconds. This information is inserted into the special metrics expression, and the expression is then attached to the corresponding event by a script. After this the expression visualizes the information automatically according to its notational context and the information stored in it. Furthermore, we are also monitoring how the user interprets the measure structure. We assume that the user will accentuate the first events of the measures. The appropriate accentuation affects the overall score by giving extra credit to the student, i.e., the absence of the accentuation would not lower the score. Finally, an overall grade is provided. The score is given as a percentage. A perfectly executed reading would result in a grade of 100%.

5 Application

PWGL provides users with several ways of organizing content: The PWGL presentation mode can be used to turn any PWGL patch into an application, and DBL (Document Builder Language) is a comprehensive application/document builder tool. The purpose of these tools is to allow for the preparation of documents that can be used for different purposes. One of the main uses is to provide rich and interactive documentation for PWGL and its user libraries. Furthermore, they provide teachers with a means to realize educational material.

Here, since our application is realized as a normal PWGL patch, the final application will be implemented with the help of the PWGL presentation mode. The presentation mode allows for the simplification and reorganization of a PWGL patch so that the pertinent components of it can be brought into the foreground without destructively changing the patch.

Figure 5 shows the complete patch implementing the generative part, the user input (the score), and the part for checking and evaluating the user input. In (1) we define the rhythmic database. The database was generated before-hand and the code responsible for it is not shown here. The constraints rules for the generative part described earlier are given in (2). They are connected to the constraints solver box (3) which applies the rules to the database to generate an exercise. Furthermore, in this patch, we also use a set of so-called heuristic rules (4). For demonstration purposes we give preference to rhythmic figures that do not contain any rests. The use of heuristic rules allows for one more

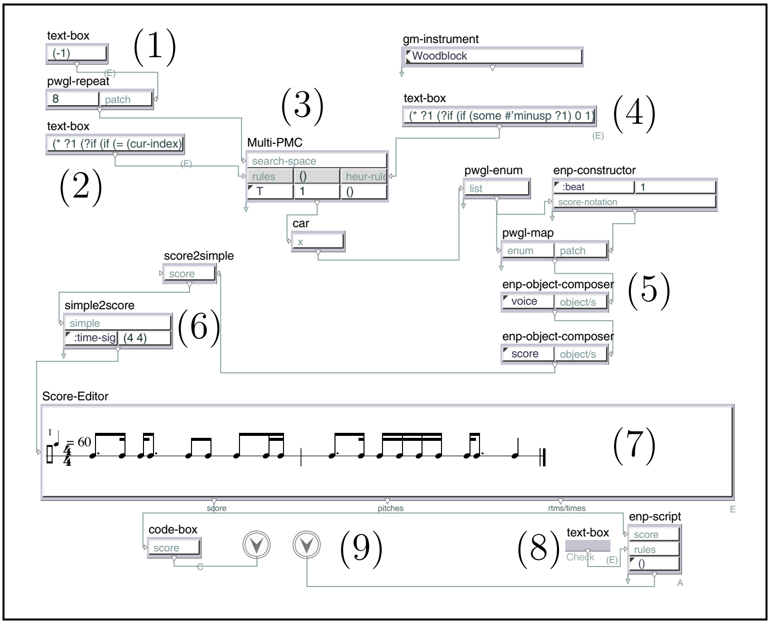


Fig. 5. The rhythm reading application implemented with the help of PWGL. The patch shown here describes the application's functionality.

Fig. 6. The user interface of the rhythm reading application. The same patch as in Figure 5 is displayed here using the 'presentation' mode of PWGL. The analysis of the reading performed by the user is displayed below the music.

level of control in terms of the degree of difficulty. The train of rhythmic figures, given as output by the constraints solver, is converted into a score object and consequently also quantized (5-6). The quantization step is effectively a rebaring step, as the rhythmic material is already in a 'quantized' form. We are using the KSQuant[12] library to apply a certain measure structure. The result, as presented to the users, is shown in the score editor (7). The validation of the user input is done with the help of the ENP-script box (8) using a set of rules. The two buttons shown in (9) are operated by the user to trigger the reading and also optionally to trigger the playback of the score.

Figure 6 shows the final application as the user would see it. Creating this view is a straightforward process. The teacher preparing the application switches to the 'presentation' mode and modifies the view as he/she sees fit, for example by hiding and displacing boxes. In our case, for instance, the end-user does not need to see the parts of the patch that implement the program logic, such as the generation of the rhythms and the checking of the assignment. Therefore, these components can be hidden from the user. Furthermore, the position and size of the remaining boxes is adjusted to create a view suitable for the task at hand. Here, the user input area is made bigger as it is of central importance.

The user interface is also relatively straightforward. Reading is initiated when the user hits the 'Start' button. After this the system waits for the first event. The reading is finished once there are enough events and/or enough time has passed after the last event.

6 Conclusions

In this paper we discuss the potential of PWGL within the framework of Computer Assisted Instruction. PWGL already exhibits several interesting features that would make it ideal for different kinds of pedagogical applications ranging from elementary music tuition to advanced studies in composition, music theory and acoustics. PWGL is unique in offering a combination of sophisticated notation and signal processing capabilities while being a full-fledged programming environment.

Here, we concentrate on the rhythm reading exercise which is one of the fundamental skills required for music students. We discuss the various components of a rhythm reading exercise application and provide an example of a reference implementation. The application is offered freely, in the hope that it may inspire and instruct the community on how to build pedagogical applications in PWGL. It may also be directly useful to teachers and students.

One of our long-term tasks would involve incorporating known teaching techniques, such as the Leitner method (also known as spaced repetition), as a part of our tuition system. Furthermore, issues involving the usability of the software from both the teachers and the students point of view should be investigated. We are also receptive to criticism and suggestions from the community of professional music educators in order to further enhance the capabilities of PWGL.

Acknowledgments. The work of Mika Kuuskankare is supported by the Academy of Finland project 137619.

References

1. Aycock, J.: A brief history of just-in-time. ACM Computing Surveys 35(2), 97–113 (2003)
2. Hofstetter, F.T.: Computer-based recognition of perceptual patterns in harmonic dictation exercises. Journal of Research in Music Education 26(2), 111–119 (1978)
3. Hofstetter, F.T.: Computer-based recognition of perceptual patterns and learning styles in rhythmic dictation exercises. Journal of Research in Music Education 29(4), 265–277 (1981)
4. Kuuskankare, M., Laurson, M.: ENP Expression Designer - a Visual Tool for Creating User Definable Expressions. In: International Computer Music Conference, Barcelona, Spain, pp. 307–310 (2005)
5. Kuuskankare, M., Laurson, M.: Expressive Notation Package. Computer Music Journal 30(4), 67–79 (2006)
6. Kuuskankare, M., Laurson, M.: Survey of music analysis and visualization tools in PWGL. In: Proceedings of International Computer Music Conference, pp. 372–375 (2008)
7. Laurson, M., Kuuskankare, M.: Extensible Constraint Syntax Through Score Accessors. In: Journées d'Informatique Musicale, Paris, France, pp. 27–32 (2005)
8. Laurson, M., Kuuskankare, M., Norilo, V.: An Overview of PWGL, a Visual Programming Environment for Music. Computer Music Journal 33(1), 19–31 (2009)
9. Norilo, V.: Introducing Kronos - A Novel Approach to Signal Processing Languages. In: Neumann, F., Lazzarini, V. (eds.) Proceedings of the Linux Audio Conference, pp. 9–16. NUIM, Maynooth (2011)
10. Percival, G., Anders, T., Tzanetakis, G.: Generating Targeted Rhythmic Exercises for Music Students with Constraint Satisfaction Programming. In: Proceedings of the 2008 International Computer Music Conference, August 24-29. International Computer Music Association, Belfast (2008)
11. Raptis, S., Askenfelt, A., Fober, D., Chalamandaris, A., Schoonderwaldt, E., Letz, S., Baxevanis, A., Hansen, K.F., Orlarey, Y.: IMUTUS – an effective practicing environment for music tuition. In: Proceedings of International Computer Music Conference, pp. 383–386 (2005)
12. Sprotte, K., Laurson, M., Kuuskankare, M.: KSQuant - complex score manipulation in PWGL through quantization. In: Ystad, S., Kronland-Martinet, R., Jensen, K. (eds.) CMMR 2008. LNCS, vol. 5493, pp. 253–261. Springer, Heidelberg (2009)
13. Toussaint, G.: The euclidean algorithm generates traditional musical rhythms. In: BRIDGES: Mathematical Connections in Art, Music, and Science, pp. 47–56 (2005)
14. Wright, M., Freed, A., Momeni, A.: OpenSound Control: State of the Art 2003. In: Proceedings of NIME, Montreal, pp. 153–159 (2003)

Which User Interactions Predict Levels of Expertise in Work-Integrated Learning?

Tobias Ley[1] and Barbara Kump[2]

[1] Institute of Informatics
Tallinn University, Estonia
tley@tlu.ee
[2] Department of Human Resources and Organization,
Vienna University of Applied Sciences of WKW, Austria
barbara.Kump@fh-wien.ac.at

Abstract. Predicting knowledge levels from user's implicit interactions with an adaptive system is a difficult task, particularly in learning systems that are used in the context of daily work tasks. We have collected interactions of six persons working with the adaptive work-integrated learning system APOSDLE over a period of two months to find out whether naturally occurring interactions with the system can be used to predict their level of expertise. One set of interactions is based on the tasks they performed, the other on a number of additional Knowledge Indicating Events (KIE). We find that the addition of KIE significantly improves the prediction as compared to using tasks only. Both approaches are superior to a model that uses only the frequencies of events.

Keywords: learning analytics, user modeling, knowledge indicating events.

1 Introduction

Learner models [1] are at the core of adaptive learning systems, as they enable a system to adapt to individual learning needs. A variety of learner characteristics (knowledge, interest, learning style, etc.) may be represented in a learner model. Within this paper we concentrate on the learner's knowledge state. To make sure that the learning system can adapt to the knowledge of its users, continuous maintenance of the learner model is necessary.

In our previous work, we have followed two alternative paths for updating the user model which both relied on implicit information derived from natural user behavior when interacting with the system: First, an approach which is based on Competence-based Knowledge Space Theory (CbKST), and second, an approach which is based on Knowledge Indicating Events (KIE).

We have suggested Competence-based Knowledge Space Theory (CbKST) as a means to derive a user's knowledge state from tracking the user's working tasks he or she has been performing as part of his or her normal job duties [2]. The user model in this case simply adds all knowledge and skills assumed to be necessary for performing well in a particular task. While CbKST has been successfully applied in educa-

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 178–190, 2013.

tional settings, certain limitations exist that make it difficult to transfer this approach to implicit user modeling in a workplace setting. There, it is difficult to observe successful task performance (implicitly) from which available knowledge and skills could be derived. At the same time, our experience has shown that users use these systems much more flexibly than simply obtaining scaffolding for performing a task. Rather, they browse or search topics, obtain learning hints, contact others or are contacted by others for help etc. [3]. As a consequence, a user model which is only based on tasks performed does not make use of all available information.

As a second approach, we have therefore suggested KIE as a means for non-invasively diagnosing user knowledge in an adaptive work-integrated learning system [4][5]. The assumption is that all behavior of a user with the system potentially can give evidence of whether a user possesses knowledge and skills. For example, an assumption could be that a person who clicks on the 'help'-button for a concept has little knowledge of this concept. While this may not be the case for each behavior, collecting enough evidence may in the end lead to a more accurate user model than if only task performance is considered.

In this paper, we compare an approach which is based on CbKST (i.e., only on tasks performed) with an approach based on KIE where a more extensive set of actions is considered. As collecting realistic user interactions in the workplace is extremely difficult and has to our knowledge not been realized, we perform an explorative study in which the work-integrated learning system APOSDLE [3] was used by six users for a period of two months. We use these two approaches to find out whether user interactions can be used to predict self-appraisal of expertise. We also compare the model to peer-assessment and to a simple frequency-based approach as a baseline.

2 Traditional Ways of Knowledge Diagnosis

In the context of adaptive learning systems, such as Adaptive Educational Hypermedia (AEH [6]), and Intelligent Tutoring Systems (ITS [7]), assessment and user model maintenance with regard to user knowledge has been typically based on *testing* for knowledge and skill. In ITS, the emphasis was placed on procedural problem solving skills. From successful or unsuccessful solutions of a problem, the skill level of a learner was deduced, and hence curriculum sequencing or learning prompts were adapted (for a recent review see [8]). The underlying models that were used in these systems include a cognitive model of the learning domain [9] [10], Constraint-based Models [11], Competence-based Knowledge Space Theory [12] or Bayesian Nets [13]. The general idea in all of these approaches has been that by tracing observable solution behavior, inferences on the underlying knowledge and skills can be made. The process of inferring knowledge in one concept from diagnosing knowledge in another concept has been called *knowledge propagation* [1] or *knowledge update* [14].

While ITS have traditionally been more concerned with testing and practicing of skills, AEH have been more concerned with teaching conceptual knowledge. As a result, the user interaction with AEH is quite different from that in ITS: In AEH, learners consume learning material by reading texts, receiving pictures or listening to audio or video materials. Learning is supposed to be triggered by these processes and is seen to be an internal process that cannot be observed directly. For this reason, AEH attempt to guide learners through the large quantities of contents in an optimal way by recommending suitable contents or sequences of materials.

The most straightforward way of diagnosing user knowledge in an AEH system is simply asking learners questions, from which their knowledge can be deduced. This happens in the form of quizzes or tests. Usually, this form is seen as the ‘most effective way’ of diagnosis [15]. There has been a realization, however, that it is not always feasible to ask questions, and thus other methods have been used. Visiting pages seems to be the most common one, which assumes that a visit indicates a certain pages has been read and understood at least to a certain degree. This measure of a user’s knowledge based on a page visit has been implemented, for example, in AHA! [16].

More recently, research into ITS and AEH has grown together, giving rise to Adaptive and Intelligent Web-based Educational Systems (AIWBES, Brusilovsky & Peylo, 2003). In the context of AIWBES, some steps have been made in the direction of non-invasive diagnosis of user knowledge. Brusilovsky & Millan [1] use the term *evidence-bearing event* to refer to a user activity that reveals that the user has knowledge about a certain domain element. These events can be “an answer to a test item, solution of a problem, teacher’s opinion, the number of Web pages relevant to the element K that have been visited, etc.” (p.26). This description of an evidence-bearing event includes the possibility of diagnosing knowledge implicitly, from a user visiting a page, as used in ElmArt [17] and LS-Plan [15], for instance. Similar approaches to assessing user knowledge from different sources of evidence have been proposed with the Cumulate server of KnowledgeTree [18] and Personis [19].

As stated above, the clear separation between ITS and AHS is gradually dissolving. Nowadays, integrated learning environments (termed, e.g. virtual learning environments or virtual classrooms) certainly contain elements of both ITS and AHS. While a learning environment may contain exercises, it could also contain learning materials that would allow a student to learn certain contents, or interactive exercises or simulations that would allow practicing certain skills. Most importantly, all modern learning environments also contain communication and collaboration functionalities which allow learners to get into contact with fellow learners, tutors or teachers. A large array of these functionalities allows for collaborative learning. Despite these developments, utilizing implicit measures for diagnosing knowledge plays a subordinate role in AIWBES. Instead, in most of these systems, testing still plays a predominant role, and other implicit forms of assessment are usually seen as less effective, less reliable and therefore as generally inferior.

Increasingly, learning environments seek to support learners not in an artificially created learning environment, but embed learning in other natural activities. Examples are game-based environments [20], or work-based learning environments that should support people in performing real tasks they are working on [23]. These systems blur the boundaries between a learning system and other types of natural activities, like natural work tasks.

3 Implicit User Modelling from Naturally Occuring Events during Work Tasks

In the context of workplace learning, where learning happens mostly in non-formal arrangements, testing for knowledge and skills (e.g. with knowledge quizzes) is not a realistic option for multiple reasons [2]. Rather, it is assumed that in workplace practices, knowledge and skills are being applied in the context of normal work tasks. Therefore, more implicit methods of knowledge diagnosis are needed.

Several implicit approaches have been proposed. The ADAPTS system [25], for example, is used for adaptive help in complex technical maintenance tasks. It combines a domain model (different components of the maintained system), a task model (a hierarchical structure of the maintenance tasks) and a user model which tracks knowledge and experience of each technician with these components and the tasks. For each task, there is a mapping to all domain elements involved in the task. The user model uses information on which tasks were successfully performed and thereby diagnoses the knowledge of a person in the tasks and components. With this system, however, no longer field study has been conducted to collect naturally occurring interactions.

One system that implicitly infers the users' IT skills, more specifically their Microsoft Word literacy, from logs is the system OWL, an acronym for 'Organisation-Wide Learning' [21]. OWL is a recommender system for learning that logs Word commands of each user such as 'print', 'copy', 'paste', etc. The commands of different users are pooled and for each single user it is decided whether he or she over- or under-uses a certain command. Under-used commands are then recommended to the users in order to improve their Microsoft Word literacy. The appeal of this approach work lies in the operationalisation of skills: The authors diagnose the user's MS Word literacy by observing his or her interaction with MS Word – practically no inference is needed for this diagnosis. This is in contrast with approaches that try to diagnose skills which are by far more difficult to observe, such as a user's knowledge in the area of 'inference statistical analyses', or 'programming'.

Other systems that give advice in the task context are recommendation systems for software development [22]. These build a user model on a per item basis and therefore do not diagnose any knowledge or generalizable skill. They observe which methods the user has already used and refrain from recommending these items again when the methods have been used for a number of times.

In the context of our own work in the area of work-integrated learning, we have proposed an approach for inferring employee competencies from past task performance in knowledge work [2]. The user model, in this case, assumes a mapping between the working tasks of a domain on the one hand, and the knowledge and skills needed to perform these tasks on the other. CbKST is then used to infer availability of knowledge and skills when performance of certain tasks is observed [23]. Because dependencies exist in the learning domain between certain knowledge and skills in the sense of prerequisites, inferences can exploit these dependencies to make the update more efficient.

Recently, and as an attempt to extend this approach, we have suggested the idea of Knowledge Indicating Events (KIE) as a means for non-invasively diagnosing user knowledge [4]. Conceptually, the idea of KIE is in line with evidence-centered assessment design as suggested by [26]. In a nutshell, KIE are naturally occurring user actions (e.g., selecting a link, accessing a learning hint) that are interpreted as evidences for a user's knowledge state. When each of these actions is connected to a concept of the domain model, then inferences can be made about the user's knowledge level about the particular domain concept.

Clearly, in order to assess the usefulness of KIE for diagnosing user knowledge in a naturalistic setting, empirical data is needed. Such empirical data requires a field study where users interact with the learning system in a naturalistic way during their daily work tasks over a longer period of time. Within the present paper we present a first field study where we statistically compare the two approaches, CbKST and KIE with regard to their success in diagnosing user knowledge. In the next section, we will introduce the field study in which the data was collected.

4 Field Study

For the field study we used APOSDLE[1], an adaptive work-integrated learning system which aims to improve knowledge worker productivity by supporting learning within everyday work tasks. Within APOSDLE, the learner model is used for ranking learning goals, recommending useful learning content, and for suggesting knowledgeable people [3][4].

4.1 APOSDLE and Its Learner Model

APOSDLE can be instantiated to various domains by creating new semantic domain models. Typically, an APOSDLE domain model consists of 100 to 150 domain concepts. The domain model also includes mappings between the domain concepts and various other elements of the system, such as work tasks, resources (document snippets), or learning paths. The APOSDLE learner model is designed as a layered overlay of the APOSDLE domain model.

To employ the KIE approach for APOSDLE, all possible user interactions with the system that could be related with one or several domain concepts were treated as potential KIE. APOSDLE tracks all user interactions with these domain model elements and from these interactions infers the knowledge state of a user for each of the concepts of the domain model based on a very simple rule-based algorithm.[2] The users receive different recommendations for a topic (different types of resources etc.), de-

[1] `www.aposdle.org`

[2] In the version of APOSDLE used in this study, the user model contained for each concept one of three knowledge levels (learner, worker, and supporter) which was also visualized in the open learner model [5]. For the present study, however, these inferences were not taken into account.

pending on the detected knowledge level. While these recommendations have been shown to improve task performance [24], in a next step, the simple rules should be replaced by a statistical approach. Therefore, the predictive power of each of the KIE that was used within APOSDLE should be investigated in a field evaluation. The question that we were posing at the outset of our explorative study was: Which user interactions predict levels of expertise in work-integrated learning?

4.2 Procedure of the Field Evaluation

The field evaluation was carried out in an innovation management company in Austria. Typical tasks in the innovation management domain are for instance, *analyzing trends in a certain industry*, or *identifying strategies of industry competitors in the market*. The APOSDLE innovation management domain comprises 144 domain concepts. Examples for domain concepts from the innovation management domain are *idea management*, *market analysis*, *patent*, or *knowledge management*.

The participants in the evaluation study were 6 employees of the company who (after being trained in the use of the system) used APOSDLE for a period of approximately 2 months in their regular work. As the interaction of the users with APOSDLE should be as realistic and natural as possible, there was no specific scenario used for the evaluation. When working on their laptops and desktop computers, the innovation consultants in the innovation management company could interact with APOSDLE at any time to search for information, and to receive learning support. For example, they could look up techniques for workshop design, or view examples of customer offers from similar previous projects. Whenever APOSDLE detected a task or topic the user was working on through automatic task detection, the user was free to view the suggestions on the task or topic provided by APOSDLE, and to explore further materials. Users were in no way forced or encouraged to use the system. So we consider the data gathered as a realistic snapshot of how users naturally interact with a work-integrated learning system in a 2-month period.

In order to establish an external criterion for the knowledge levels in each topic, values of the expertise level of a user in each topic were needed. As mentioned above, objective knowledge testing is not possible in workplace learning for a number of reasons (no tests are available and there is large resistance of experts to be tested for their expertise). We therefore drew on self and peer appraisal which is a very common form of appraisal in workplace settings. For example, it is used in a number of personnel development tools, such as 360 degree feedback instruments [27]. Furthermore, while research has shown the biased nature of self-appraisal [28], there is evidence to suggest that it can be a valid form of assessing states of knowledge in technical areas and for experienced job holders [29].

Therefore, a special type of card sorting was applied. Card sorting as a technique for software evaluation has also been used, for instance by Wild et al. [30]. For self- and peer-assessment, each of the participants was provided with a set of 144 cards, one for each topic in the domain. In a first self-assessment trial, the participant was asked to sort each of the cards (topics) into one of five categories: (i) “I am rather inexperienced in this topic”, (ii) “I am neither very inexperienced nor very experienced in this topic”, (iii) “I am very experienced in this topic”, (iv) “This is not my area of work”, and (v) “I do not understand the topic”.

In a second run, two peer-assessments were gathered for each participant. Therefore, each participant was asked to sort the same cards for two of their colleagues. In addition to the five categories in the self-assessment, a sixth category, (vi) “I cannot decide on the level of expertise of my colleague in this topic”, was available for peer-assessment. That way, for each participant, we obtained one complete self-assessment and two complete peer-assessments for each topic (domain model element) in the innovation domain.

4.3 Results

4.3.1 Knowledge Indicating Events Collected in the Study

We define KIE as traceable naturally occurring interactions of a user with a learning environment that potentially allow inferences on the user’s knowledge and skills. In our view, KIE can provide both positive evidence for knowledge, and negative evidence of knowledge, that is, evidence for the absence of knowledge. Examples of KIE include contacting a person about a topic, or annotating a document with a topic.

APOSDLE provides a variety of functionality, including the adaptive presentation of learning goals, the recommendation of knowledgeable colleagues, or the recommendation of documents and text passages. For the KIE approach, we identified all possible user interactions with APOSDLE that could possibly indicate user knowledge. All possible user interactions the system that could be related with one or several domain concepts were treated as potential KIE The rationale for not pre-selecting any type of interaction was the exploratory approach in this study: We were interested in the predictive power of the different types of KIE. The list of KIE can be seen from the first column in Table 1.

The KIE *perform topic* indicates that APOSDLE automatically detected a topic on which the user was working on (e.g., the user was creating a presentation where he or she used the term “innovation management”). Another KIE is called *select learning goal – topic*. This KIE implies that a person has clicked on an item in a list of learning goals which lead to the display of further information for the topic mentioned in the learning goal. The KIE *view resource* indicates that a user viewed a resource (e.g. a report, presentation, image or video) annotated with a certain topic. The KIE *edit annotation* means that a user annotated a resource with a certain topic, or modified an existing annotation. The KIE *perform task* means that a user carries out a task which requires knowledge about certain topics. This task-topic relationship is represented in the APOSDLE domain model. The KIE *contact peers* means that a user contacted another APOSDLE user via APOSDLE to communicate about some topic within the APOSDLE domain model. A KIE *select learning goal task* exists that indicates that a person has clicked on an item in a list of learning goals which were presented for a task the user was working on. The KIE *get learning hints* means that a user requested additional ‘learning hints’ (questions, exercises) for a topic at hand. The KIE *create learning path* indicates that a topic occurred in a intentionally created ‘learning path’ (a list of learning goals which are arranged in a didactically beneficial sequence). Finally, the KIE *being contacted* means that a user was contacted by another user via APOSDLE about a certain topic.

Table 1 gives an overview over the system usage of each participant in the testing phase of APOSDLE in terms of the frequency of how often each KIE was applied by each of the participants (P1 to P6). On average, a user applied 303.83 KIE (s =288.81) during the evaluation phase. P1 applied 875 KIE; hence, she used the system at most and rather frequently. P4, P5, and P6 showed moderate usage in terms of number of KIE. P2 and P3 showed a low usage of the system with less than 120 KIE.

What becomes obvious already from the table is the unbalanced distribution of KIE types: the most frequently applied KIE were performing a topic (855), selecting a learning goal for a topic (399), viewing a resource (180), editing an annotation (176), and performing a task (149). All other KIE occurred much less frequently (some even only sporadically).

Table 1. Frequency of knowledge indicating events (KIE) for each of 6 participants (P1-P6)

KIE	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*	*Total*
Perform topic	375	59	36	118	143	124	855
Select learning goal-topic	158	16	29	66	65	65	399
View resource	116	3	7	8	36	10	180
Edit annotation	143	3	2	2	18	8	176
Perform task	68	3	26	23	12	17	149
Contact peers	2	6	7	2	--	8	25
Select learning goal-task	5	5	3	1	1	--	15
Get learning hints	--	--	5	--	--	6	11
Create new learning path	5	--	--	1	1	2	9
Being contacted	3	--	--	1	--	--	4
Total KIE	875	95	115	222	276	240	1823

4.3.2 Predictive Models

In order to predict the level of expertise, a model was needed that could predict as a criterion a user's level of expertise (the user's self-assessment of being a beginner, advanced or an expert) for each of the 144 domain topics using different types of predictors (e.g. tasks only vs. all tracked events). We chose to use a multiple linear regression model as a prediction model which fits a linear regression line through the data by minimizing the squared difference between actual category values (self-assessed expertise) and prediction. The overall regression coefficient (R) gives an overall indication of the fit of the model, while the standardized regression weights (beta) give an indication of the weight each predictor carries in the prediction.

We were only concerned in this analysis with predicting the level of expertise (beginner, advanced, expert) for each topic and each user. We therefore filtered the data for all cases where we had both a valid expertise rating from the user (leaving out the ones were the user indicated they did not understand the topic or it was not in his/her working domain) and for which there was at least one event that had been tracked by APOSDLE. The reason for the latter was that we did not want the data to be too skewed by the different levels of engagement of the different users. So the model should have a realistic chance to make predictions from the collected data. Analyses were performed on a "per person" and "per topic" level so that each data point corresponds to one person and topic combination. The dataset we used contained 340 such data points of 864 possible ones (6 users times 144 topics). All analyses reported below were performed on this dataset of 340 data points.

We are giving below the results of four different predictive models. All models use the user's self-assessment as a criterion. In a baseline model, we use the two assessments by colleagues as predictors. Although the input data for the human baseline model differs from the other models (colleagues assessed users on a 3-point scale, while the KIE models use all gathered activities as predictors), we still use the colleague assessment model as a baseline to compare all models based on KIE to human judgment. Note that it is likely that humans form judgments about other's expertise in a similar way as APOSDLE, namely by gathering evidence from behaviors and forming a composite judgment.

We then employed three KIE models to compare their predictive potential. First, we employed a model that took only the performed tasks into account. This is in line with CbKST which considers for each task all the skills assigned to that task and assumes the user possesses these skills if the task is performed successfully. In our case, we do not know about the success of task performance, but we assume that the higher the number of detected or selected tasks, the higher the probability that the user can in fact perform the task successfully, and, hence, possess the skills assigned to the task. We call this model *Task KIE model*.

The second KIE model used as predictors the frequency of all events, not only the performed tasks. As can be seen from Table 1, the amount of information this model used is significantly larger. While the KIE task model drew only on 149 events, the *Full KIE model* drew on 1823 events.

Finally, we also computed a model which did not take into account the type of event that was performed for a concept, but rather drew only on total frequency of events for a concept (*Frequency KIE model*). So, for example, if a user had performed one task related to the concept, had looked at two resources and had created one learning path for the topic then the Full KIE model would take these different events into account by using all of them as predictors. The Frequency KIE model instead would only use one predictor for the topic, namely total number of events, in this case 4.

To summarize, the following 4 models were computed

- *Human judgment model* (baseline model): 2 Predictors: two expertise ratings from colleagues
- *Task KIE model*: 1 Predictor: Frequency of KIE relating to the performance of a task
- *Full KIE model*: 9 Predictors: Frequency of KIE for all 9 event types.
- *Frequency KIE model*: 1 Predictor: Frequency of all KIE for a topic, regardless of the type.

4.3.3 Comparison of Linear Regression Models

Table 2 gives an overview of the outcomes of the linear regression to predict self-assessed knowledge levels using the four models. The baseline model (human judgment) reveals a coefficient of R=0.405 and is significant with F(2,534)=53.198, p<.001. These results indicate that, taken together, the two peer assessments predicted the self-assessment to a moderate degree, and predictions were clearly better than chance. The beta weights indicate that both peer judgments contribute significant amounts of variance to the model: $beta_{peer1}$=0.340 (t=8.268, p<0.001), $beta_{peer2}$=0.139 (t=3.370, p=0.001). Results for these peer assessments are in line with correlation coefficients usually obtained in studies that compare self- and peer assessments, for instance [31] and [32] (see below in the discussion section).

As can be seen, two of the KIE models are significant (Task KIE and Full KIE model) while the Frequency KIE model is not. On the one hand, this gives general support to the two approaches we have pursued. As CbKST suggests, looking at the tasks performed at the workplace, it is possible to predict to some extent expertise needed to perform these tasks. Also our second approach to track activities with APOSDLE was successful to predict expertise better than chance.

Moreover, if these two models (tasks only vs. all events) are compared to each other, then the addition of explained variance (the difference in R^2 of .053) turns out to be significant ($F_{change}(8,330)=2.382$, p=0.017). This indicates that the Full KIE model is superior to the Task KIE model in that it adds to the prediction some additional variance.

Table 2. Results of four linear regression models to predict self-assessment of expertise for 144 domain topics

	No. of Predictors	*R*	*F*	*p*
Human judgment model	2	0.405	53.198	<0.001
Task KIE model	1	0.147	7.514	0.006
Full KIE model	9	0.274	2.979	0.002
Frequency KIE model	1	0.071	1.714	0.191

4.3.4 Split Half Validation of Beta Weights

Next, we look at the beta weights of the predictors in the Full KIE model. Because beta weights as such should not be interpreted as signifying their general importance in predicting the criterion, a split half validation was performed. The data set was split into half by random assignment of all cases. Then the linear regression model was calculated for each half and beta weights were derived. This procedure should ensure that the most important beta weights could be identified. Table 3 shows the most important three predictors in each of six iterations and the sign of the beta weight. The signs of the weights in the table show the direction of the prediction where a positive sign indicates that the predictor indicated increased expertise. All signs are in an anticipated direction, except for “Get a Learning Hint” which indicated increased expertise, rather than less.

The table confirms that performing a task and creating a learning path were the two dominant predictors as they appeared among the top three in each of the cross validation datasets. The third predictor varies across the datasets indicating that additional variance cannot be explained in a stable manner.

Table 3. Three strongest predictors and the sign of their beta weight in six split half cross validation samples

Perform Task (+)	Create Learning path (-)	Create Learning path (-)	Create Learning path (-)	Create Learning path (-)	Perform Task (+)
Create Learning path (-)	Perform Topic (+)	Perform Task (+)	Perform Task (+)	Perform Task (+)	Create Learning path (-)
View Resource (+)	Perform Task (+)	Get Learning Hint (+)	View Resource (+)	Select Learning Goal (-)	Get Learning Hint (+)

5 Discussion and Conclusion

We conclude from the results that it is possible to predict to a certain extent self-assessed expertise in a domain by considering events tracked by a work-integrated learning system in a natural use of the system at the workplace. We note further that an approach, based on CbKST which predicts expertise solely from tasks selected is already successful. If we add further events which are assumed to indicate knowledge (KIE), then the prediction can be significantly improved.

The correlation coefficients derived from judgments of peers (one peer, R=0.340; two peers R=0.405) are in line with studies that compare self- and peer assessment in workplace settings. For example, two meta-analyses find correlation coefficients between self- and peer ratings of ρ=0.36 [31] and r=0.37 [32]. Considering this as an upper baseline, we consider the performance of the Full KIE model (R=0.274) as a rather promising result, considering the fact that the KIE model did only have the opportunity to track events for 2 months, and there was a great variation in how frequently APOSDLE was used by the participants.

Our study also clearly shows that it makes sense to look at qualitative differences in the events rather than just count frequencies. There is a good amount of evidence in the behavioral and educational sciences that supports this finding: When compared to novices, experts in a domain don't just do more of the same, but instead their behavior and thinking is qualitatively different than that of novices. Comparing the Task and Full KIE models to a model based on tracking frequencies of events only, clearly shows the superiority of taking into account different types of events. This is also supported by different signs of the predictors when the beta weights were more closely analyzed.

It should be mentioned that we do not intend to suggest that a linear regression model is the most suitable statistical model to predict expertise. Instead, the purpose of the present paper was merely to compare the predictive power of different KIE models in a situation of real workplace behavior. Linear regression models have a wide applicability and they have been shown to be fairly robust, and this is why they have been our choice in this case. We did employ some alternative regression models (like logistic regression) and found the general conclusions to be the same. We are fairly certain that the predictive accuracy could be improved by employing more sophisticated machine learning models.

Finally, it should be mentioned that the generalizability of these findings needs further research in other domains. Especially, the beta weights clearly need further validation. It is quite likely that the importance of each of the types of events for predicting expertise will vary significantly between domains, or even from person to person. If enough data was available, it might be worthwhile to build a predictive model for each person which from a user modeling point of view may even improve the performance of the algorithm.

Acknowledgements. APOSDLE (www.aposdle.org) has been partially funded under grant 027023 in the IST work programme of the European Community.

References

[1] Brusilovsky, P., Millán, E.: User Models for Adaptive Hypermedia and Adaptive Educational Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) Adaptive Web 2007. LNCS, vol. 4321, pp. 3–53. Springer, Heidelberg (2007)

[2] Ley, T., Kump, B., Albert, D.: A methodology for eliciting, modelling, and evaluating expert knowledge for an adaptive work-integrated learning system. International Journal of Human-Computer Studies 68, 185–208 (2010)

[3] Lindstaedt, S., Kump, B., Beham, G., Pammer, V., Ley, T., Dotan, A., de Hoog, R.: Providing varying degrees of guidance for work-integrated learning. In: Wolpers, M., Kirschner, P.A., Scheffel, M., Lindstaedt, S., Dimitrova, V. (eds.) EC-TEL 2010. LNCS, vol. 6383, pp. 213–228. Springer, Heidelberg (2010)

[4] Lindstaedt, S.N., Beham, G., Kump, B., Ley, T.: Getting to know your user - Unobtrusive user model maintenance within work-integrated learning environments. In: Cress, U., Dimitrova, V., Specht, M. (eds.) EC-TEL 2009. LNCS, vol. 5794, pp. 73–87. Springer, Heidelberg (2009)

[5] Kump, B., Seifert, C., Beham, G., Lindstaedt, S.N., Ley, T.: Seeing what the system thinks you know. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, LAK 2012, pp. 153–157. ACM Press, New York (2012)

[6] Brusilovsky, P., Kobsa, A., Vassileva, J. (eds.): Adaptive hypertext and hypermedia. Kluwer Academic Publishers, Dordrecht (1998)

[7] Poulson, M.C., Richardson, J.J. (eds.): Foundations of intelligent tutoring systems. Lawrence Erlbaum Associates, Hillsdale (1988)

[8] Desmarais, M.C., Baker, R.S.J.: A review of recent advances in learner and skill modeling in intelligent learning environments. User Modeling and User-adapted Interaction 22, 9–38 (2011)

[9] Anderson, J.R., Boyle, C.F., Corbett, A.T., Lewis, M.W.: Cognitive modeling and intelligent tutoring. Artificial Intelligence 42, 7–49 (1990)

[10] Ritter, S., Anderson, J.R., Koedinger, K.R., Corbett, A.T.: Cognitive tutor: Applied research in mathematics education. Psychonomic Bulletin & Review 14, 249–255 (2007)

[11] Mitrovic, A.: Fifteen years of constraint-based tutors: What we have achieved and where we are going. User Modeling and User-Adapted Interaction 22, 39–72 (2011)

[12] Heller, J., Steiner, C., Hockemeyer, C., Albert, D.: Competence-based knowledge structures for personalised learning. International Journal on E-Learning 5, 75–88 (2006)

[13] Conati, C., Gertner, A.S., VanLehn, K.: Using Bayesian Networks to Manage Uncertainty in Student Modeling. User Modeling and User-Adapted Interaction 12, 371–417 (2002)

[14] De Bra, P., Aroyo, L., Cristea, A.: Adaptive web-based educational hypermedia. In: Web Dynamics: Adapting to Change in Content, Size, Topology and Use, pp. 387–410 (2004)

[15] Limongelli, C., Sciarrone, F., Temperini, M., Vaste, G.: Adaptive learning with the LS-Plan dystem: A field evaluation. IEEE Transactions on Learning Technologies 2, 203–215 (2009)
[16] De Bra, P., Smits, D., Stash, N.: The Design of AHA? In: Proceedings of the ACM Conference on Hypertext and Hypermedia, Odense, Denmark, p. 133 (2006)
[17] Brusilovsky, P., Schwarz, E., Weber, G.: ELM-ART: An intelligent tutoring system on World Wide Web. In: Frasson, C., Gauthier, G., Lesgold, A. (eds.) ITS 1996. LNCS, vol. 1086, pp. 261–269. Springer, Heidelberg (1996)
[18] Brusilovsky, P.: KnowledgeTree: A distributed architecture for adaptive E-Learning. In: WWW 2004, New York, USA, May 17-22, pp. 104–113 (2004)
[19] Kay, J., Kummerfeld, B., Lauder, P.: Personis: A server for user models. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) AH 2002. LNCS, vol. 2347, pp. 203–212. Springer, Heidelberg (2002)
[20] Augustin, T., Hockemeyer, C., Kickmeier-Rust, M.D., Albert, D.: Individualized skill assessment in digital learning games: Basic definitions and mathematical formalism. IEEE Transactions on Learning Technologies 4, 138–148 (2011)
[21] Linton, F., Schaefer, H.-P.: Recommender systems for learning: Building user and expert models through long-term observation of application use. User Modeling and UserAdapted Interaction 10, 181–208 (2000)
[22] Happel, H.-J., Maalej, W.: Potentials and challenges of recommendation systems for software development. In: Proceedings of the 2008 International Workshop on Recommendation Systems for Software Engineering, RSSE 2008, vol. 11. ACM Press, New York (2008)
[23] Ley, T., Ulbrich, A., Scheir, P., Lindstaedt, S.N., Kump, B., Albert, D.: Modelling competencies for supporting work-integrated learning in knowledge work. Journal of Knowledge Management 12, 31–47 (2008)
[24] Ley, T., Kump, B., Gerdenitsch, C.: Scaffolding Self-directed Learning with Personalized Learning Goal Recommendations. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 75–86. Springer, Heidelberg (2010)
[25] Brusilovsky, P., Cooper, D.W.: Domain, task, and user models for an adaptive hypermedia performance support system. In: Proceedings of the 7th International Conference on Intelligent User Interfaces, IUI 2002, vol. 23. ACM Press, New York (2002)
[26] Mislevy, R.J., Riconscente, M.M.: Evidence-centered assessment design. In: Downing, S.M., Haladyna, T.M. (eds.) Handbook of Test Development, pp. 61–90. Lawrence Erlbaum Associates, Mahwah (2006)
[27] Lepsinger, R., Lucia, A.D.: The art and science of 360 degree feedback. John Wiley & Sons (2009)
[28] Hoffman, C., Nathan, B., Holden, L.: A comparison of validation criteria: Objective versus subjective performance measures and self- versus supervisor ratings. Personnel Psychology 44, 601–619 (1991)
[29] Muellerbuchof, R., Zehrt, P.: Vergleich subjektiver und objektiver Messverfahren für die Bestimmung von Methodenkompetenz - am Beispiel der Kompetenzmessung bei technischem Fachpersonal. Zeitschrift für Arbeits- und Organisationspsychologie 48, 132–138 (2004)
[30] Wild, F., Haley, D., Bülow, K.: Using latent-semantic analysis and network analysis for monitoring conceptual development. Journal for Language Technology and Computational Linguistics 26, 9–21 (2011)
[31] Harris, M.M., Schaubroeck, J.: A meta-analysis of self-supervisor, self-peer, and peersupervisor ratings. Personell Psychology 41, 43–62 (1988)
[32] Conway, J.M., Huffcutt, A.I.: Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. Human Performance 10, 331–360 (1997)

Motivating Students or Teachers?

Challenges for a Successful Implementation of Online-Learning in Industry-Related Vocational Training

Nils Malzahn, Tina Ganster, Nicole Sträfling, Nicole Krämer, and H. Ulrich Hoppe

University of Duisburg-Essen, Duisburg, Germany
{malzahn,hoppe}@collide.info,
(tina.ganster,nicole.straefling,nicole.kraemer}@uni-due.de

Abstract. In this paper we present our findings from the FoodWeb2.0 project about success indicators and restraints while implementing Web2.0 based learning processes. We conducted two courses with the same content with two different target groups from the German food industry: one with regular employees and one with teachers of dedicated education facilities. Comparing the performance of the two courses by triangulating methods from Social Network Analysis and quantitative and qualitative surveys, we identified indicators for the successful implementation and differences in the motivation of learners and teachers. These findings illustrate the need for strategies involving and motivating teachers when introducing modern learning methods and tools within the food industry and other related branches.

Keywords: Learning analytics, Teaching techniques and strategies for online learning, Learner motivation and engagement, Vocational training, CSCL, Web2.0, Collaborative Learning, SNA.

1 Introduction

The food industry in Germany is characterized by a high amount of workers without or with only a low level of formal qualification. While these workers are easily found and taught to perform the simple and often physically exhaustive tasks, there is a lack of employees with a higher qualification (e.g. skilled workers), who are able to use and control the complex machines and processes of the food production industry. Thus, the human resource managers try to train some of the lowly qualified to a higher qualification level to close the gap. This is not an easy task since these people often have a migration background and therefore language problems and/or they are not very motivated to learn because of various reasons, e.g. education is not an asset to them or their work is so exhausting that they are not ready to learn after work. The project FoodWeb2.0 (funded by the German Ministry of Research and Education) aims at training the employees of the German food industry using two basic strategies: Motivating employees for vocational training and performing education in collaborative, blended learning using Web2.0 technologies. We propose this approach,

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 191–204, 2013.

because we think that Web2.0 tools offer appropriate support for collaborative learning activities that support the development of valuable competences and skills with respect to their workplace requirements.

Furthermore, Web2.0 tools implicitly contain functionalities that provide gratifications through media usage [1] by addressing relevant user needs. These needs comprise several informational, entertainment or social aspects [2,3] that are also relevant for learning processes and motivation (e.g. social comparison, gamification and social feedback). Thus, the use of Web2.0 technologies (e.g. profile pages, rating systems) may increase the learners' motivation for learning.

During the FoodWeb2.0 project we conducted several courses on the FoodWeb2.0 platform using a theory guided approach to motivation and learning. In the following sections we want to present our findings about success indicators and restraints, when using a blended learning approach with respect to motivation, successful learning and perceived vs. observable collaboration using Social Network Analysis methods in combination with standardized interviews.

2 Collaborative Learning Using a Web2.0 Platform

There are a lot of studies that deal with the effect of collaborative learning on learning outcomes and success. Tannenberg [4] found that collaborative learning is perceived as less tiring than other learning methods when learning process flows. Kulick [5] found that collaborative learning increases the intrinsic motivation for the learning topic. Therefore, it is likely that the willingness to engage oneself in the learning topic after having finished the course and the motivation to participate in further trainings will increase. Since Baird and White [6] observed that collaborative learning encourages the learners to be responsible for their own learning, it is expected that the participants of a formal learning process will take part in other non-formal learning processes as well.

Many adult training courses are characterized by heterogeneous groups of learners. As collaborative learning promotes peer learning [7], a collaborative learning approach helps to even the different skill levels during a course because students learn from each other and weaker student may profit from the knowledge of skilled learners within the course.

Additionally, collaborative learning encourages the acquisition of desired social [8] and leadership skills [9]. Thus, collaborative learning is capable to improve the learning success of a wide variety of courses for diverse target groups.

Another problem is the short time spent at the training facility. Thus, the examples shown or the example processes experienced by the participants of such a course cannot represent the complexity of the real processes at the participants' workplaces. Hence, the participants experience difficulties in applying the newly acquired knowledge to their everyday work. Davidson [10] showed that collaborative learning may lead to a better simulation to the workplace situation, because it allows for a division of labor and therefore to work on more complex examples.

Last but not least, collaborative learning is known [11] to bring forward stable groups of learners or even learning communities. This helps to sustain a learning process even after a course finishes, because trust built during the initial collaborative learning process is transferred to subsequent situations at the workplace. Thus, workplace problems are discussed within the learning group, stimulating informal learning.

In summary the usage of collaborative learning methods is expected to have positive effects on learning outcomes and its respective results. For a start it is irrelevant if collaborative learning is used in presence learning or online learning as some of the findings are observed in traditional classroom settings without computer support. Although the authors are not aware of any study that proves that Computer Supported Collaborative Learning (CSCL) increases the positive effects of collaborative learning, Lethinen [12] at least shows that these effects are also achieved in computer supported settings.

The philosophy of Web2.0, including user generated content, collaboration on artifacts, as well as mutual sharing of artifacts and knowledge, fits well with basic idea of collaborative learning. Thus, using Web2.0 tools to leverage the advantages of CSCL in the food industry seems promising. Hence, we set up a Web2.0 learning platform. According to Kerres [13], a Web2.0 learning platform should be a flexible environment providing services that enable the learner to make active and passive use of the tools of the Internet. This implies that the learning platform is not only the provider of content and tools, but also a guide post to external content and tools. To enable the learners to organize their learning process, a learning platform should allow for the creation and management of a personalized learning environment (PLE), which ideally adapts itself to the individual learning context [13]. King's [14] classification of learning that distinguishes instruction (goal-oriented, formal instruction), explorative learning (goal-oriented, informal) and serendipity learning (without a goal, exploring, informal) fits well with Kerres' [13] definition of a Web2.0 learning platform. Traditional e-learning1.0 is almost completely based on instructional learning [13]. Content is provided to the learners to be consumed by them. Explorative learning [14] demands that the learners navigate through the learning content on their own. That means that the learners should be able to move through the content outside of the limits of a single course or lesson to still their learning needs. In this case serendipity learning is a piggyback effect, since the typical Web2.0 link collections (either tagged or otherwise indexed) often provide the opportunity of discovering an interesting bit of information by chance.

The FoodWeb2.0 project provides a platform for all three types of learning as well as the opportunity to create a PLE. Every contribution on the platform may be tagged, commented and rated depending on the configuration by the creator of an artifact, by a teacher (for respective courses) or the site administrators. There are several possibilities to aggregate collections of contributions depending on their affiliation to a course, a topic or the artifact's type (article, blog, wiki, video, file, image, forum etc.). Furthermore, the basic platform system (Liferay Portal) allows for the configuration of a personal area that may be used as a PLE. Site-wide provided templates for reasonable tool combinations help beginners to create their own PLE: Thus, the FoodWeb2.0 platform provides a complete environment for Web2.0 learning.

To investigate the effect on the learners' motivation we set up a course on Web2.0 tools and good practices, because the participating industry partners wanted to train their employees on this topic.

3 Case Study – A Course about Web2.0

The course teaches basic knowledge about using Web2.0 tools and good practices (privacy, intellectual property, „friendship“, etc.) on the Internet in general. The basic structure of every lecture consists of introductory media (text, video, images) provided by a teacher, followed by task descriptions asking the participants to create contributions (images, videos, text). These contributions vary from comments and forum posts with respect to “central questions” in the beginning to collaborative work in small groups within a wiki (Jigsaw design [15]) in the later phases of the course. Furthermore, the course employs WebQuests [16] to motivate the learners to explore important topics of the course and share their acquired knowledge with each other.

Thus, the participants are getting to know (and to use) the Web2.0 tools step-by-step while learning about important topics with respect to the Web2.0 in general.

As the course participants were distributed all over Germany, the course was conducted completely over the Internet, framed by two video conferences: a kick-off meeting, where the basic platform functions (logging in, navigating through the course content etc.) were explained and a feedback meeting after the final exam, where the participants were asked to give feedback and tell about their experiences with this type of course, which was new for them.

There was no general video conference in between, but the teachers offered individual online meetings to aid the participants in case of technical or other problems. There was just one participant who asked for additional help. The course officially concluded with a final exam. If the exam was passed, a certificate of successful participation was issued.

3.1 Participants’ Course Evaluation

In order to identify success indicators and restraints for Web2.0 courses, we conducted two runs of the course with different target groups. The first run (A, 14 learners) was conducted with regular employees of the German food industry. The second run (B, 13 learners) was conducted with persons teaching in current trainings for the German food industry. Thus, we are able to compare the behavior and performance of employees and the teachers in this industry branch when confronted with Web2.0 learning. Before and after each course, a survey was conducted. In the following sections a descriptive comparison of the results of the survey is triangulated with the results of a Social Network Analysis based on the users’ actions logged by the portal system (cf. **Fig.**1) to get an insight into the motivational differences of the participants of the two course runs.

Demographics. Participants in both courses who completed the survey (N_A=6, N_B=8) covered a broad age range from 20-59 years. Participants in course A were less educated (1 person had a college degree) than those in the course B (5 of whom had a college degree).

Fig. 1. Snapshot of the FoodWeb2.0 platform and its integrated analysis cockpit

Prior Knowledge. Participants' self-assessment of their knowledge and use of Web2.0 was measured using seven items (four-point Likert scale; 1= do not agree; 4 = totally agree). Results show that participants in both courses had similar prerequisites (see **Fig.**2). Since both courses consisted of volunteers we may assume that the participants of both courses were already interested in Web2.0 in general.

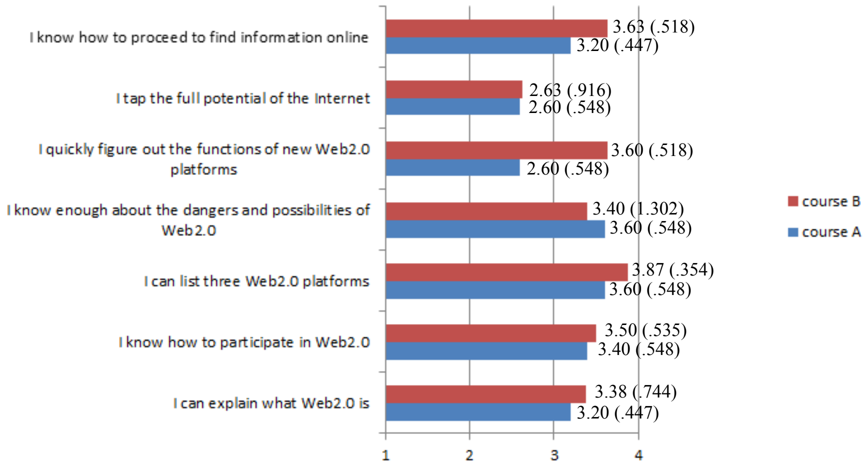


Fig. 2. Participants' prior knowledge (self-assessment)

Performance. The overall performance in both courses in terms of number of contributions to the course and successful final exams is almost the same. The slight difference in the overall contribution amounts (see **Fig.3**) can be explained by the different amount of participants (A: 14, B:13, without course instructors). The amount of static, instructor generated content (article) is the same in both courses.

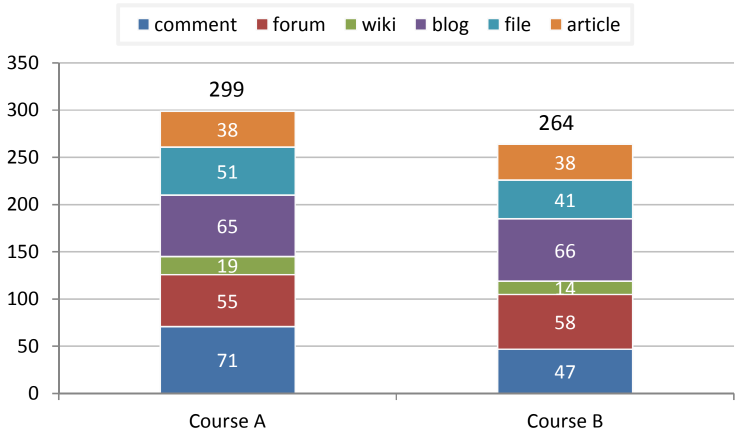


Fig. 3. Distribution of artifacts among the different courses (A above B)

General Course Satisfaction. The survey included five items (rated on five-point Likert scales; 1=not at all satisfied, 5=absolutely satisfied) to assess participants' satisfaction with different aspects of the course. The results presented in **Fig.4** indicate that participants of course A were generally more content.

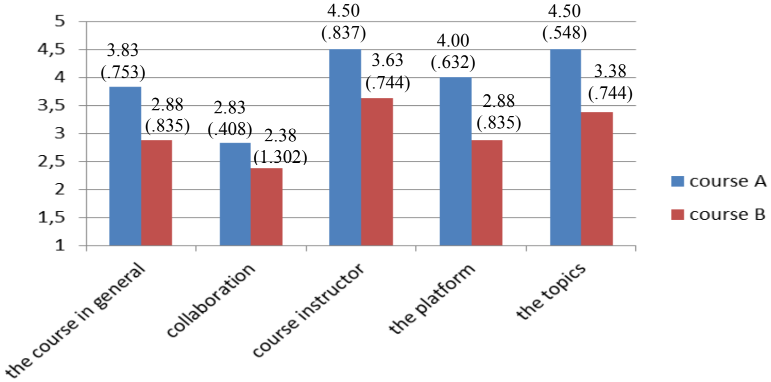


Fig. 4. Satisfaction with different aspects of both courses

Additional open-ended questions were included that asked participants to indicate what they liked most and least about the course.

With regard to positive issues, participants in both courses mentioned similar aspects: the combination of theoretical aspects and practical applications (mentioned 2x in course A and 3x in course B) and sufficient support from course instructors (mentioned 1x in each course). Additionally, two participants from course A addressed the issue that the course allowed for flexibility regarding working hours.

The aspects participants liked least about the courses differed between groups: In both courses, people addressed problems in collaboration (3x in each course). But while in course A these problems were assigned to differences in participants' states of knowledge regarding Web2.0, participants in course B criticized the type of assignments as a potential cause. In line with that, participants in course B found the assignments unrewarding and had the feeling that they did not allow for knowledge transfer. Additionally, participants in course B addressed technical problems and usability as negative aspects about the course.

Learning. Participants in both courses were asked (open-ended) which aspects they found most helpful for learning. Participants in course A addressed practical application of learned content (1x), contributions by other participants (1x) and exploring topics (1x). Participants in course B mentioned the provided texts (1x), additional links (1x) and RSS feeds (1x).

A quantitative assessment (items rated on four-point Likert scales; 1=not helpful; 4= very helpful) confirms the impression that for participants in course A, practical application and exploring topics on their own was helpful for learning while participants in course B preferred learning with the information provided by course instructors. Particularly with regard to the helpfulness of exploring topics on their own, participants in course A and B held different opinions (**Fig.5**).

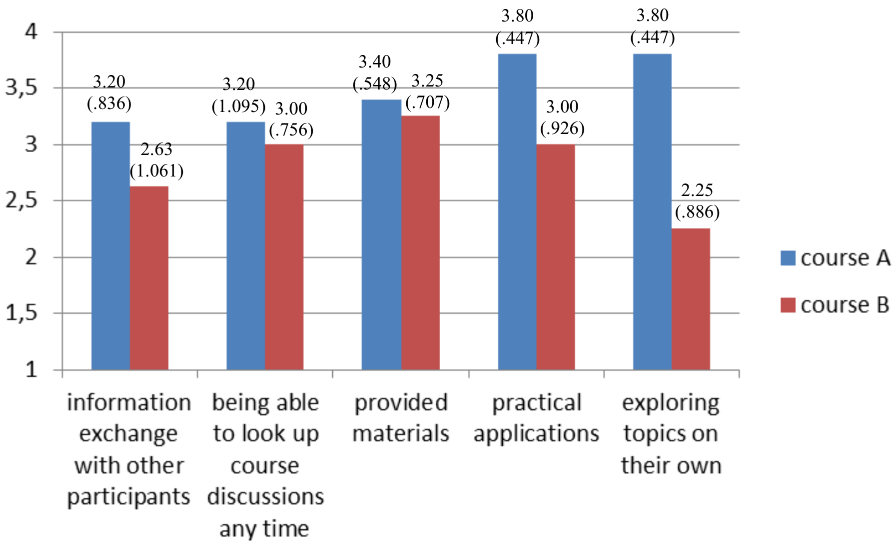


Fig. 5. Aspects perceived to be helpful for learning

Collaboration. Several items were included to assess participants' perception of collaboration (rated on a four-point Likert scale; 1=do not agree at all; 4= totally agree).

Results (see Fig.6) show that overall, participants' impression of collaboration in course A was more positive than in course B. It needs to be mentioned that standard deviations in course B were higher than in course A, indicating a less consistent evaluation across participants in course B.

Fig. 6. Perception of collaboration

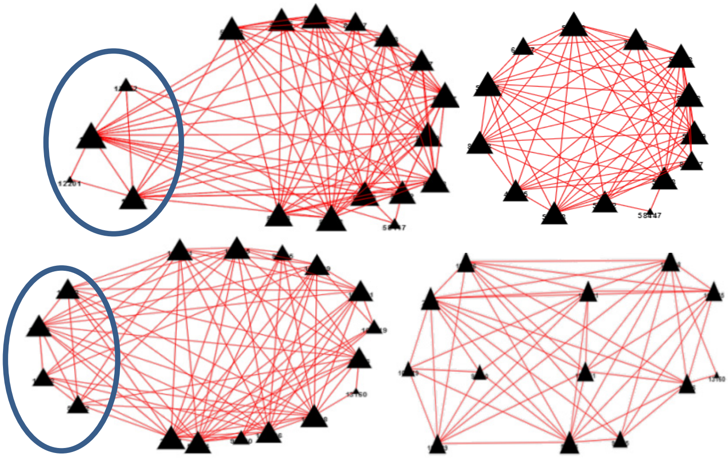


Fig. 7. Collaboration between participants with (left) and without teachers (right) (A above B) as logged by the system

Since the surveys just present the *perceived* collaboration, we complemented the survey with a Social Network Analysis based on the *observed* artifact interaction of the participants. **Fig.7** shows the results of a degree based analysis of the interaction. The first row shows the interaction in course A. The second row shows the results from course B. The left column shows the resulting collaboration network including the teachers (in the blue ellipsis). The right column shows the same network without the interaction with the course instructors. The size of the triangles representing the members of the course indicates the degree the respective node. A higher degree results in a bigger triangle. The size increases logarithmically to emphasize the order of difference.

The SNA results show that the perceived incorporation of course B cannot be proved by the actual interaction. In both courses the participants interact directly with each other. The networks do not fall apart when the course instructors are removed from the network. Nonetheless, the density of network A (ρ=0,71) is higher than the density of network B (ρ=0,64).

3.2 Discussion

Overall, participants in course B were less satisfied with the course. Quantitative and qualitative data suggests that problems of participants in course B at least partially were grounded in the type of assignments and the resulting group work. They found the approach of exploring topics on their own (in groups) less helpful for learning than participants in group A. Further results indicate that course B would have preferred to work individually and to be provided only with the necessary information to work on the assignments. Consequently, group B contributed fewer comments on the work of others (see **Fig.7**). High standard deviations in course B (particularly with regard to collaboration) indicate that not necessarily all participants shared the same negative view, but that some participants were more involved in group work than others as can be seen from **Fig.7**. While the magnitude of interaction per participant (size of triangle) in course A was almost the same for all participants (**Fig.7**; upper right network), the size (based on a logarithmic scale) of the triangles in course B indicates bigger differences between the respective participants. This correlates to Swan's findings [17] which states: „Students reported high levels of interaction with their classmates also reported higher levels of satisfaction and higher levels of learning from course."

Since the general topic of the course is Web2.0, the course instructors used a lot of the Web2.0 tools offered by the FoodWeb2.0 platform. In turn they also asked the participants to use these tools giving clear instructions on how to do so. The participants used all of the tools as demanded and expected to solve the given tasks, but the use of forums and blogs were preferred over wikis and videos. Additionally, in course A the possibility to link one's profile with other members of the course was used a lot. This might indicate a bigger sense of belonging to a group, which indicates a better satisfaction of the motivational aspect of "need to belong"[18].

Although there is some criticism with respect to the provided information within the course, the average uptake of the course's contents is good. The solutions to the given tasks were generally good and the effort for the participants was as expected, i.e. some of the participants needed more time than the average participant and others needed less. Thus, it can be stated that the course was successful from a teacher's point of view. The notion that the participants spent additional time on the provided information because they valued this information supports the hypothesis that Web2.0 learning is capable of initiating informal learning processes and leads to (motivates) an increased interest in the covered topics.

Since there are differences in the perception and acceptance of the course design, we conducted an additional study with course instructors (teachers) of the institutions participating in FoodWeb2.0. We asked 19 teachers if they have used other learning management system (e.g. Moodle) than FoodWeb2.0 in their courses, what they think about collaborative learning and used a standardized personality survey [19] to get information about their attitude towards online teaching and innovative teaching.

With respect to collaborative learning some of the teachers (n=3; M=4,5; Likert-scale of 5) stated that they use collaborative learning more often than the others (n=16; M=3,0).

Two of the teachers think that it is difficult to document the results of collaborative work. This opinion is not shared by those, who already have experience with online learning (n=10). In both groups there are teachers who think that the teacher stays more in control of the learning process, if the course is teacher- centered.

In general all of the teachers think that a mix of collaborative learning and teacher presentation is the best way of teaching.

Looking at the result of the personality survey, we found that all of the teachers are generally open for new experiences.

Although the general attitude towards Web2.0 learning, collaborative learning and online learning as well as prior experience with online learning does not indicate considerable differences between the teachers on the platform, the success of the courses with respect to motivation of students and implementation of Web2.0 learning (or just collaborative learning) varies.

While four of the teachers have successfully implemented a Web2.0 blended learning course and additional 6 are using Moodle for blended learning without collaboration, nine teachers were not able to conduct an online course with considerable participation of the students.

Thus, there is no correlation between open mindedness and collaborative learning or between the stated attitude towards collaborative learning and the successful implementation of Web2.0 courses.

4 Indicators for Successful Courses

As the survey results do not indicate that the success of a Web2.0 course depends on the teachers' attitude towards the Internet and collaborative learning, we have tried to identify best practices for successful Web2.0 courses. From these best practices we derived some indicators for successful courses that may be used to intervene, if a course is becoming unsatisfactory.

The presented case studies on the FoodWeb2.0 platform show that the use of pedagogical patterns [20] are generally a good indicator for successful courses and can be adapted to fit Web2.0 learning platforms and motivate interaction with the platform and each other. Examples from other training organizations show that ignoring the motivational needs of the participants or the usage of pedagogical anti-patterns (e.g. unclear task description, free choice of (Web2.0) tools, omitting introduction of the virtual classroom or the tools to be used) lead to unsuccessful courses (see **Fig. 8**), especially in blended learning courses. This confirms that Web2.0 technology as a sole motivational aspect of a course is not sufficient, but the question must be raised why some course instructors are more successful than others while adapting their curricula to a (Web2.0) learning platform.

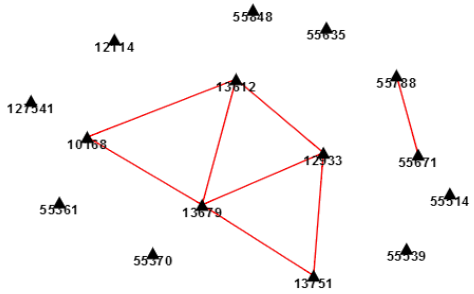


Fig. 8. Collaboration network of an unsatisfactory course: almost only the five teachers collaborate

The use of WebQuests in adult trainings, especially in Web2.0 based learning, has received a positive feedback by the training institutions participating in the Food-Web2.0 project. The method is even able to ease the cold start problem, because the participants are asked to contribute additional content to initial set of provided material. Thus, a growing set of material is created with each course run. Even inexperienced participants are able to provide valuable contributions. Furthermore, the introduction of new tools should be linked to content-rich tasks to motivate the use of the tools and force the participants to deal with it.

Another basic indicator for a successful course is the distribution of artifacts within a course. The more artifacts are generated in a collaborative tool like wikis, forums or blogs the better the interaction between the participants. If there is a significantly high amount of "static" content like "html content", the course is most often not successful. **Fig. 9** shows a selection of courses run on the Foodweb2.0 platform. While two of the courses (the Web2.0 courses from above) made use of almost all available Web2.0 tools, other courses like no. 63138 relied almost only on forums, but there are many entries made by the students. Course 29081 and 130966 are not considered successful. The acceptance by the students was low and therefore they did not contribute to the course.

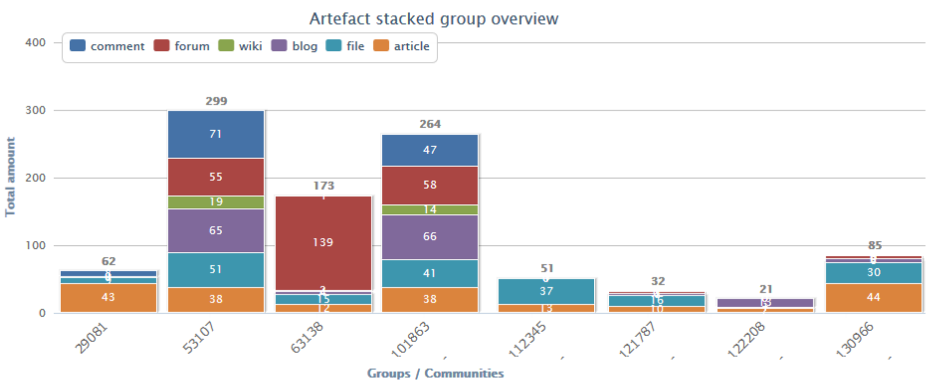


Fig. 9. Distribution of artifacts among several courses

Although the current data set is not sufficient to give a final answer, there are some indicators. The participants of course A seem to be more open-minded with respect to the Web2.0 learning. On the one hand they had less experience (self-assessment) than the members of course B. On the other hand the participants of course B (mainly teachers) seem to be more skeptical concerning Web2.0 and the Internet in general. This is supported by other studies both conducted by us and others as well [21]. Additionally, the teachers have their own experiences and set ways of teaching that have to change. Thus, the negative statement that individual work would have been more effective than the group work is almost as old as collaborative learning itself [22], which shows that motivating teachers to use a learning platform while introducing the platform may be even more important than motivating learners. We will elaborate on this in the concluding section.

5 Conclusion and Implications for Change Management

In this paper we identified basic success indicators for a blended learning in vocational training and showed that there is a considerable difference in the perception of blended learning (esp. Web2.0 learning) between vocational trainers and regular employees of the German Food industry. This study together with other studies conducted within the project suggest that this cannot be explained just by open-mindedness, attitude toward collaborative learning or expertise with internet technologies (esp. Web2.0 technologies), because we could not find a clear correlation or even causality with these indicators. What we found in other studies was that the bigger institutions have more problems than the smaller ones. There seem to be organizational issues that have to be taken into account. A regional study [21] supports this hypothesis. This study [21] also found that educational institutions are reluctant to introduce Internet technologies into their courses because of missing media competence and Web2.0 experience at the decision maker level.

Thus, it is not only a problem of the teachers, but the teachers do not get the proper motivation and free time to adapt their courses to the new technologies. Consequently, teachers who already have some experience with these technologies are less concerned about the dangers of Web2.0 technologies, but they have a realistic estimation about the needed effort for the initial setup of an online course. Teachers with almost no experience are afraid of the technological dangers and overestimate the effort, but both groups would try the new technologies if they had time.

If this time is organized by the decision makers we may observe that the teachers recognize the added value of a teaching directed at digital collaboration within the course and co-operation with peer teachers. The perceived effort for the teachers and the students even decreases.

It is observed that the inexperienced teachers lose their reservation towards Internet technologies quickly and contribute in creative discussions on how to use a particular tool adequately during their teachings.

> *„My expectations and hopes in this technology regarding didactics and education and everything, were satisfied. There was learning growth observable among the students. It works!"* (Teacher of another course on FoodWeb2.0).

Some of the more advanced teachers even employ some kind of KAIZEN-like [23] quality control and continuing improvement process by standardizing their teaching processes and teaching material to enable a re-use of different modules in different courses.

In spite of the success with motivated teachers in an adequate organizational context ((paid) time for discussion, (paid) time for adaption, (paid) time to get to know the tools, a slot for tool introduction in curriculum of a course), a teacher who is not convinced of the use of technology will result in students that do not (want to) use the tools or the platform.

> *„One disadvantage that I see is that it all costs a lot of time. I had the luxury of being exempt from my usual duties to become acquainted with the platform."* (Teacher of another course on FoodWeb2.0)

Thus, it is important to focus on a proper motivation of the teachers at first. It seems reasonable to begin with a selected group of pilot teachers that are especially cared for. The transition from traditional classroom lectures to blended learning lectures should be as smooth and easy as possible.

> *„Motivating colleagues here is not that difficult. I told them to give their materials to me so that I can upload them onto the platform for them to get started"* (Teacher of another course on FoodWeb2.0)

Thus, we provided tools that are capable of emulating PowerPoint presentations to re-use the available material as a start. Afterwards new tools and learning processes are introduced with the pilot teachers step-by-step. These pilot teachers will show their success to their peers and the decision makers, so that a subtle process of convincing other teachers is triggered.

Nonetheless, the greater goal has to induce a technological awareness as well as an innovation awareness (w.r.t. teaching) to allow for further innovation (not only for teaching). For this goal the pilot teachers are important for the motivation of their peers, because they ground the innovations into the social and professional context of the institution with their personal credibility. This implies that the teachers participating in innovating projects are part of the permanent staff and not only freelancers, who are hired to fulfill project requirements or to hold courses like it is wide-spread in food industry training organizations, to ensure a sustainable impact on the current situations and to avoid the social dilemma of not sharing knowledge with other teachers to have a competitive advantage for further employments.

References

1. Katz, L., Blumler, J.G., Gurevitch, M.: Uses and Gratifications Research. The Public Opinion Quarterly 37(4), 509–523 (1974); HAAS u.a. 2007
2. Raacke, J., Bonds-Raacke, J.: MySpace and Facebook: Identifying Dimensions of Uses and Gratifications for Friend Networking Sites. Individual Differences Research 8, 27–33 (2010)
3. Haferkamp, N., Krämer, N.C.: Creating a Digital Self. Impression Management and Impression Formation on Social Networking Sites. In: Drotner, K., Schröder, K.C. (eds.) Digital Content Creation: Perceptions, Practices & Perspectives, pp. 129–146. Peter Lang, New York (2010)

4. Tannenberg, J.: Using Cooperative Learning in the Undergraduate Computer Science Classroom. In: Proceedings of the Midwest Small College Computing Conference (1995)
5. Kulik, J.A., Kulik, C.C.: College Teaching. In: Peterson, P., Walberg, H. (eds.) Research in Teaching: Concepts, Findings and Implications. McCutcheon Publishing (1979)
6. Baird, J.R., White, R.T.: Improving learning through enhanced metacognition: A classroom study. In: Annual Meeting of the American Educational Research Association, New Orleans (1984)
7. Cohen, E.G.: Restructuring the classroom: Conditions for productive small groups. Review of Educational Research 64(1), 1–35 (1994)
8. Breen, P.: 76 Career-related liberal arts skills. AAHE Bulletin 34(2) (1981)
9. Johnson, R.T., Johnson, D.W.: Using cooperative learning in math. In: Davidson, N. (ed.) Cooperative Learning in Mathematics: A Handbook for Teachers. Addison-Wesley (1990)
10. Davidson, N. (ed.): Cooperative Learning in Mathematics: A Handbook for Teachers. Addison-Wesley (1990)
11. Tinto, V.: Enhancing learning via community. NEA Higher Education Journal 6(1), 53–54 (1997)
12. Lehtinen, E., Hakkarainen, K., Lipponen, L., Rahikainen, M., Muukkonen, H.: Computer Supported Collaborative Learning: A Review. (1999), `http://www.comlab.hut.fi/opetus/205/etatehtava1.pdf` (access March 31, 2013)
13. Kerres, M.: Potenziale von Web 2.0 nutzen. In: Hohenstein, A., Wilbers, K. (eds.) Handbuch E-Learning. DWD Verlag, München (2006)
14. King, F.B., Young, M.F., Drivere-Richmond, K., Schrader, P.G.: Defining Distance Learning and Distance Education. Educational Technology Review 9(1), 1–14 (2001)
15. Aronson, E., Patnoe, S.: Cooperation in the classroom: The jigsaw method. Longman, New York (1997)
16. Dodge, B.: FOCUS: Five rules for writing great WebQuests. Learning & Leading with Technology 28(58), 6–9 (2001)
17. Swan, K.: Virtual interaction: Design factors affecting student satisfaction and perceived learning in asynchronous online courses. Distance Education 22(2), 306–331 (2001)
18. Baumeister, R.F., Leary, M.R.: The need to belong: Desire for interpersonal attachments as a fundamental human motivation. Psychological Bulletin 117(3), 497–529 (1995)
19. Rammstedt, B., John, O.P.: Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 41(1), 203–212 (2007)
20. Bergin, J., Eckstein, J., Manns, M.L., Sharp, H. (eds.): Patterns for Active Learning (2004)
21. SVEA: Zusammenfassung der regionalen Bedarfsanalyse – Der Einsatz von Web 2.0 in der Berufs- und Erwachsenenbilddung in Baden-Württemberg (2010), `http://www.svea-project.eu/fileadmin/_svea/downloads/Zusammenfassung_der_regionalen_Bedarfsanalyse_Baden-Wuerttemberg.pdf` (access: March 31, 2012)
22. Green, N., Green, K.: Kooperatives Lernen im Klassenraum und Kollegium - Das Trainingsbuch. Seelze (2005)
23. Imai, M.: Kaizen: The Key to Japan's Competitive Success. McGraw-Hill, Irwin (1986)

Shared Annotations: The Social Side of Exam Preparation

Andrea Mazzei[1], Jan Blom[1], Louis Gomez[2], and Pierre Dillenbourg[1]

[1] CHILI, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
{andrea.mazzei,jan.blom,pierre.dillenbourg}@epfl.ch
[2] Learning Sciences, University of California, Los Angeles CA 90095-1521, USA
lmgomez@ucla.edu

Abstract. In this paper we show how the deployment of a lightweight note sharing system can restore the antique social vocation of annotations in the classroom. The system was designed for the classroom context and evaluated through a longitudinal study lasting for an academic semester and involving 20 participants, enrolled in a Master-level course in computer science. Three key findings emerged. First, the tool spontaneously became an integral part of the classroom learning practices. Students took and shared annotations during the lectures and used them as complementary preparation material for the exam. Second, a correlation was observed between the annotation browsing time and the final exam grade. Third, a social bias emerged in favor of accessing one's own and friends' annotations. Based on the results, we discuss potential design implications for the system.

Keywords: Shared Annotations, Enhanced Learning, Social Bias.

1 Introduction and Motivation

The social vocation of annotations slipped away after the advent of the printing press. Originally the manuscripts were annotated only by relevant academic authors. The produced glosses were shared to facilitate the scholarly need for elucidation and reinterpretation of the obscure passages and terminology of the texts. Nowadays students take notes individually, and engage occasionally in the spontaneous practice of note sharing through informal meetings, email or other forms of electronic communications. Communication features explicitly aimed at facilitating such note sharing practices could have benefits on the individual learning, group cohesion and team building. In the last semi-centennial we have assisted to the emerging evidence that note taking, in its different forms and fashions, serves both as an *encoding* and an *external storage function* [30]. Taking notes can trigger a deeper processing of information [5] and therefore increase learning and recall performance [12]. It has been also observed that annotations, in the form of *visual arguments*, can encode textual information in spatial forms that require a lower understanding effort [31] and help users in establishing relationships between the content and their mental constructions [17]. These theoretical foundations and empirical observations gave us the motivation in developing *annOot*, a system for sharing and browsing annotated instructional materials. This tool was embedded within the information system of the Swiss Federal Institute of Technology, Lausanne, and used by

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 205–218, 2013.

20 students to reinforce the content associated with a Master's level computer science course. During the course, each student was provided with a tablet and a digital pen for reading and annotating the course material. The system collected both the produced annotations and the browsing history data for the entire semester. Through the analysis of these data we discuss the spontaneous adoption of *annOot* into the classroom learning practices. Then we point out salient navigation patterns and speculate on the social impact that shared annotations have within a community of learners.

2 Related Literature

2.1 The Vanishing Gap between Paper and Digital Annotations

The mid 90s are known for the explosive digitalization of the publishing industry. At that time some of the widely used commercial authoring tools came to include features for in-context digital annotating. However, many authors observed and questioned the surprisingly poor adoption of web-based forms for annotating in the educational and work contexts [10]. O'Hara and Sellen [24] formalized the advantages that paper can offer to the annotation process, compared to computer-based interfaces, which were claimed to be cumbersome and to present an obstacle for the deep reading process. They also discussed major design implications for digital reading devices based on a seamless integration of annotating and reading. In a similar fashion, many of the commercial computer-interfaced authoring tools for annotation were taken under examination in order to define a set of design recommendations for note-taking applications [25]. Later Wolfe proposed a framework for discussing annotation technologies [33]. She spanned and classified many technologies for producing external textual attachments to any digital document. She concluded by shifting the focus of the discussion towards the pen-tablet XLibris [28], which was, at the time, a novel annotation tool able to bring some of the paper affordances and digital system capabilities into the reading process. Its ergonomic design succeeded in mimicking some of the physical properties of paper annotations. A usability test reported negligible differences on the reading habits using XLibris compared to the pen-and-paper [19]. More commercial solutions are available nowadays, some of those that were experimented in the classroom context, are reported in Section 2.2.

A different approach was initiated by Wellner with the DigitalDesk [32]. He came up with the idea of using paper as interface over digital resources. A camera-projector system was used to identify and track both the paper and the user inputs and to augment them with a visual feedback. Such an interaction paradigm blurs the line between paper and digital maintaining all the physical affordances of paper and including many of the valuable strengths of digital forms and tools. Optical acquisition and character recognition constituted a viable technological base for paper-based interfaces. They allowed the integration of storing, retrieving, and manipulating capabilities into paper documents [27], and linking of on-line daily planners with templated paper agendas [13]. More recently Bonnard et al. [3] showed how paper interfaces can be easily adopted into the traditional classroom practices. The miniaturization of commercial cameras, that allowed the integration of the optical acquisition within the writing implement giving birth to the era of digital pen. A commercial implementation of such a device is

Anoto, a digital pen that incorporates a small camera on the tip. These devices allowed the emergence of many frameworks for taking digital annotations on paper texts [11] and for sharing them with teachers and students [18].

2.2 Research on Social Annotations

Digital annotation tools introduced a social component in the annotation practices. The CSCL community took the opportunity to incorporate such social forms of annotation to facilitate the teaching-learning processes. For instance, Wellner explored remote collaborative writing with Digital Desk [32]. Davis and Huttenlocher [7] reported a longitudinal study concerning CoNote; a web-based system that enabled users to access on-line course discussions via shared annotations. The system was initially deployed in the course of computer science and then became integral part of the teaching tools instantiated for that course. The authors reported anecdotal evidence on the usefulness of the tool in learning the course material and how easily the annotations were integrated within the course material. Crossen et al. [6] experimented a reader awareness system, using the XLibris tablet [28], in a high school classroom, reporting emerging needs from usability perspectives both on students and instructor sides. Another tablet-based device for note taking which was successfully experimented in classroom lectures was NotePal [8]. Marshall and Brush [20] investigated what type of annotating practices were more likely to be shared and how the transition from personal to public annotations should be mediated by mobile technology. Miura et al. [22] experimented the system AirTransNotes for augmenting in-classrooms activities using shared notes. A successful use of tablet for cooperative note-taking was also reported by Kam et al. [15]. Tablets-mediated communication was also successfully experimented in the classroom to gather students annotations as feedback of the teaching activity [1,2].

3 annOot

As reported in the literature review, both paper- and tablet-based applications can constitute valid settings for a seamless integration between in-classroom note-taking and the teaching activities. However it seems that *More Schools Embrace the iPad as a Learning Tool* [14]. For our technological setting we decided to opt for a tablet-and-pen[1] based solution because it provides ergonomic features for taking notes in symbiosis with the learning space and compatible with situated learning practices. Also it brings all the browsing, searching and storing features of mobile computers. This setting was coupled with the web-based system *annOot*, that we specifically designed, developed and deployed in the longitudinal study, reported in Section 5. *annOot* is a system providing browsing capabilities of the annotated course material. The system offers learners with an exhaustive and enhanced navigation of each other's digital notes. The users of *annOot* can take notes directly on the tablet, using commercial applications, during the lecture (as exemplified in Figure 1). Then they can revise the annotations and later decide to upload them on the *annOot* server using proprietary means of sharing. After every lecture each user can browse all the shared annotated material by using the web-based graphical user interface of the tool (see figure 2).

[1] Second generation *iPad* coupled with a *Just Mobile AluPen*

Fig. 1. Note-taking during lectures using tablets

3.1 Server Architecture

The server side embodies a database containing each single page or slide of the digital instructional material of the course. It can collect all the shared annotated documents via HTTP receiver or by polling a dedicated file sharing system. When the user decides to share an annotated document the corresponding document images are submitted to the server. The server receives and processes them, using a computer vision procedure, in order to separate the annotations from the original document. More technically each shared page/slide is first recognized using *locally likely arrangement hashing* [23]. Second, it is rectified based on the estimated planar pose [9]. Finally a subtraction procedure, resistant to moderate misalignment, is applied to the rectified annotated document image in order to extract the layer containing the annotations.

3.2 Graphical User Interface

A student can access a specific annotated slide by using *annOot*. He or she has to go through the following five steps, highlighted in the five labeled areas in Figure 2.

1. Enter the url *annoot.epfl.ch*. The address shows that the system belongs to the realm of EPFL web information systems. We decided to deploy *annOot* as EPFL web service in order to exhibit credentials trustful for the users.
2. Select the *visualization modality*. The *individual* modality allows users to visualize all the annotated material for a specific annotator. The *group* modality allows users to visualize all the annotations for a specific page. The *Moodle* button allows users to go to the page of the course containing all the annotation-free course material.
3. Choose the *lecture date*. Each button corresponds to a given lecture of the course.
4. Choose the *slide number*. Once clicked on a chosen lecture date, the user can display a specific annotated slide by clicking on the slide menu located in the bottom of the page. The menu looks like a histogram. Each bar of the histogram corresponds to a slide of the lecture material. The height of each bar quantifies the amount of annotations produced for that slide. Left and right arrow keys were also included as navigation hot keys for scrolling the slide show.

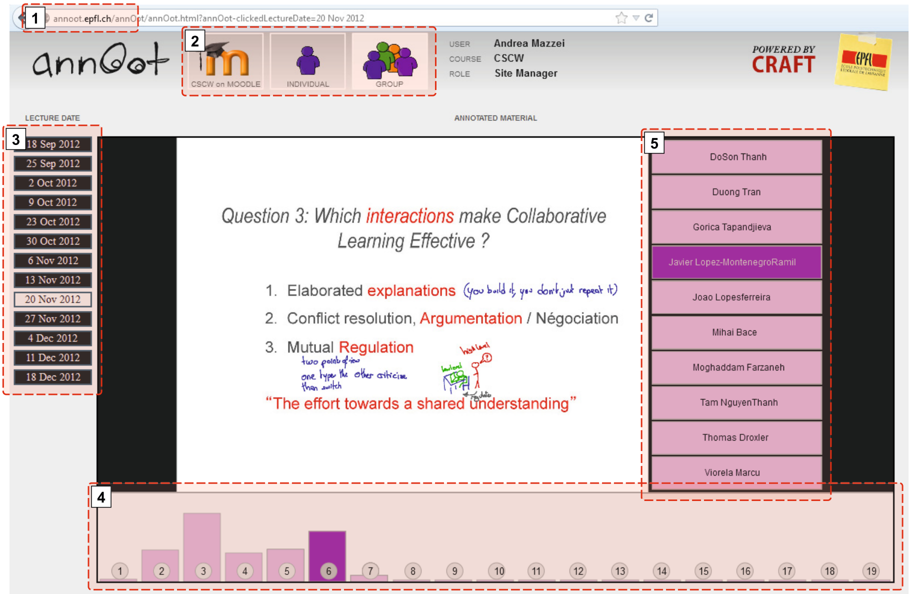


Fig. 2. *annOot* graphical user interface

5. Select the *annotator name*. The user can display the annotations of a student by clicking on his or her name in the menu displayed on the right side of the page. Through this panel a user can immediately recognize his or her own and friends' annotations. In Section 6.4 we discuss how this panel can convey a social bias affecting the user browsing behaviour.

To summarize, we see the following two modes of usage characterizing the system. First, annotations can be made to the lecture handouts when listening to the lecture. Second, the taught content can be reinforced at any given time by reading through the lecture slides and accompanying annotations. In the next two sections we state our research hypotheses concerning the usage and utility of the tool and we describe the study that we have conducted in order to assess them.

4 Research Questions and Hypotheses

Two main user practices concern *annOot*: note taking in the classroom and browsing shared annotated material after the lecture. Our very first question was centered on a potential relationship between these two user behaviours.

Q_A. How does note-taking affect the browsing behaviour of annotated slides?

– H_1. *The number of shared annotations, taken during a lecture, has a positive relationship with the time spent browsing all the annotations taken for that lecture.*

Second, we reconsidered the theoretical and empirical foundations sustaining that, taking notes [30,5,12], and reading in presence of external representations [31,17], enhance learning and retention. Therefore we wondered whether using *annOot* while taking and sharing notes and browsing shared notes would be beneficial for students comprehension.

Q_B. How does a frequent usage of the *annOot* system affect the student learning performance?

– H_2. *The number of notes taken during a lecture, being aware of sharing them later, is positively correlated with the final exam grade.*
– H_3. *The time spent browsing the shared annotations as complementary exam preparation practice is positively correlated with the final exam grade.*

Third, considering the increasing evidence that information propagation in social networks is subject to the *homophily principle* [21], we investigated whether the classroom social boundaries influence the browsing behaviour of the shared annotated material.

Q_C. How does the social influence affect the navigational flow of shared annotated material? We expect that a friendship bias might influence the way a user opens a new slide by first clicking on the name of the annotator to visualize the annotations. This could be investigated by the following formal hypotheses:

– H_4. *The empirical probability of first clicking on an annotation produced by the annotator him/herself is higher than the a priori theoretical probability.*
– H_5. *The empirical probability of first clicking on an annotation produced by a friend of the annotator is higher than the a priori theoretical probability.*
– H_6. *The empirical probability of first clicking on an annotation produced by a non-friend of the annotator is lower than the a priori theoretical probability.*

In addition, we expected a friendship bias also on the browsing time of the annotations:

– H_7. *The time spent browsing an annotation produced by a friend is higher than the time spent browsing an annotation produced by a non-friend classmate.*
– H_8. *The time spent browsing one's own annotation is lower than the time spent navigating an annotation produced by anyone else.*

5 Study Design

A longitudinal study, starting at the beginning of October 2012 and ending at the end of January 2013, was conducted in order to address the hypotheses postulated in section 4.

5.1 Participants and Instructional Material

The group of Master students, at the Swiss Federal Institute of Technology Lausanne, that enrolled in the course Computer Supported Cooperative Work in the fall semester 2012 was invited to participate in the study: 20 students accepted to participate, 4 declined their interest in the experiment, 1 dropped the course. Among the participants there were 5 females and 15 males. Their age ranged between 21 and 28 years old (M = 23.6, SD = 1.18). The course material is constituted by 13 sets of slides, counting a total of 386 slides. Complementary resources referring to books and media contents were indicated on the Moodle page of the website.

5.2 Procedure

The study was organized during the fall semester 2012. At the beginning, each participant was given a tablet and a digital pen and was instructed on how to download the course material onto the tablet and how to use it to take notes. A commercial free application for note taking was suggested, however all participants were free to choose their preferred one. Finally they were explained the necessary procedural steps for sharing notes and introduced to the main features of *annOot* browsing capabilities. All participants could keep the tablet as personal device for the entire semester; however they were asked to bring it to the classroom for the lectures. At the end of the semester each student was evaluated via formal oral examination. The examination committee was composed of two teachers of the course and an external expert in the field. The examiners elaborated 24 open-ended comprehension questions covering the content of the course material. During this examination instance, each student randomly chose two questions. The three examiners had to reach a consensus on the final grade, considering the appropriateness and completeness of the student's answers. All grades were expressed in the numeric range between 1.0 and 6.0, with a minimum passing threshold of 4.0.

5.3 Data

The collected data can be divided among three classes: (1) system usage, (2) social network; (3) mapping between lecture handout content and exam questions.

System Usage. Two types of data were recorded from the participants. First, all the annotations were collected from the students for all the course lectures, and stored in *annOot* database. Second a user tracking system was developed in order to collect data from the usage of *annOot* navigation tool. For each action performed on *annOot*, the following information were recorded: the time-stamp, the user identifier, the visualization mode, the lecture identifier, the browsed slide and selected annotator's name.

In order to shed light on the user browsing behavior of shared annotated slide an inferred variable was computed: the annotated slide browsing time (*ABT*). This variable measures how much time the user spends browsing the shared annotated slides. The variable is calculated by subtracting two sequential timestamps referring to two different browsed annotation identifiers and the same user identifier. We performed a measurement check of the *ABT* in order to filter out some noise from the data. For instance a user could forget to close the page of the browsed annotated slide creating artificially long navigation times. Therefore we first imposed a maximum threshold on the *ABT* of 5 minutes. In addition, during some lab tests we observed that the user could be tempted, by some navigational properties of the graphic user interface, to rapidly browse the annotated material without paying attention to the content of the annotations. To verify this observation we magnified the histogram of the *ABT* in a narrow range comprised between 0 and 5 seconds (see figure 3). The histogram revealed the existence of a first peak at 200 ms circa. This peak contains all the *ABT* data entries that are shorter than the average reading fixation time of 250 ms [29]. Therefore we concluded that all these short *ABT* were not representative of the meaning expressed by the variable and we filtered them out by imposing a minimum threshold of half a second.

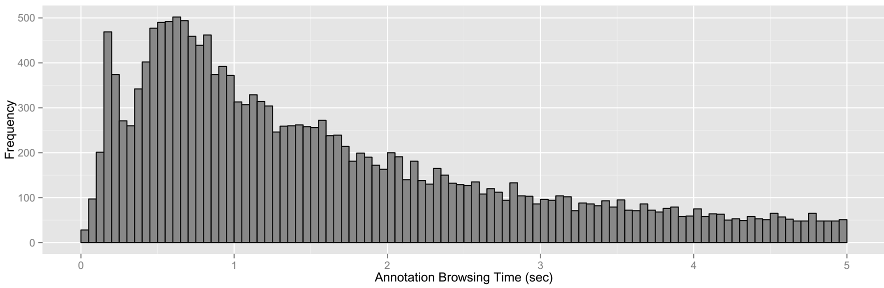


Fig. 3. Histogram of the *ABT* in the range 0 - 5 sec.

Friendship Cognitive Social Structure. As formally expressed in Section 4, a perceived friendship bias in the user browsing behavior was expected. To address this research hypothesis it was necessary to construct a data structure modeling the friendship ties in the classroom. The experimented release of *annOot* did not include any social networking capability. Therefore at the end of the course we decided to collect data on perceived friendship by using individual on-line surveys (we guaranteed anonymity conform to the data protection laws). Then we encoded the perceived friendship relationships, extracted from the on-line surveys, into *cognitive social structures*, by using means and modalities similar to the ones proposed by Krackhardt [16]. In this procedure we collected one friendship map for each respondent. This data structure allowed us to elaborate a categorical variable indicating the social provenience of a browsed annotation (*ASP*). This variable indicates whether a browsed annotation was originally produced by *ego* (the user herself), or by *friends* or by *non-friends*. To obtain an overview of the friendship network of the classroom, we generated the friendship socio-gram of the *locally aggregated structure* [16]. The data structure is a simple graph $C = (S, F)$. Each node S corresponds to a student. Each edge F corresponds to a mutually perceived friendship; for instance the edge $e = (s_i, s_j)$ is the friendship expressed by the student s_i with regards to the student s_j and vice versa.

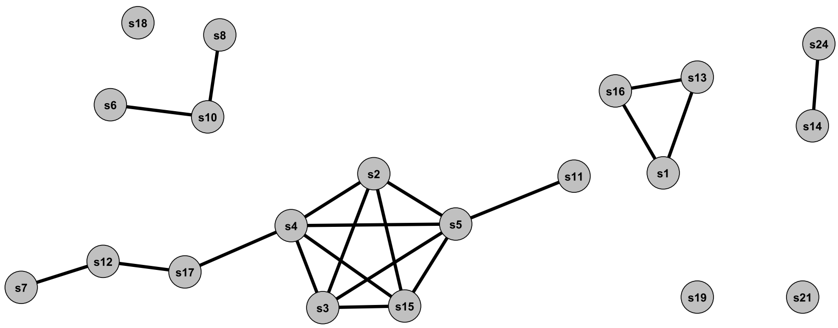


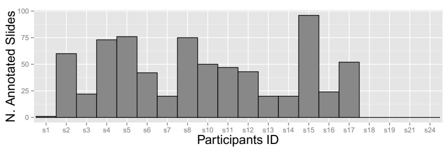
Fig. 4. Socio gram of the *locally aggregated structure* [16]

Mapping between Handout Content and Exam Questions. Chi [4] suggested a method for reducing the vulnerability of qualitative data analysis to subjective interpretations. The method applies to coding qualitative data. Inspired by this approach we elaborated a matching scheme between exam questions and instructional units. Each of the 24 exam questions was matched with the subset of lecture slides that should be read and understood in order to answer to the question itself. We modeled this question-slides dependency using a bipartite graph $G_Q = (Q, L, E)$. $Q = \{q_1, q_2, ..., q_n\}$ corresponds to the set of questions asked to the students during the examination session. $L = \{l_1, l_2, ..., l_k\}$ corresponds to the set of slides released during the course. The bipartite graph was manually constructed by one of the authors, by linking the questions in Q to the sets of lecture slides in L, such that if the understanding of the content of the set of slides l_i is necessary to answer to the question q_j, the edge $e_{i,j}$ is added to the set E. This bipartite graph allowed us to recompute all experimental features only for the partitioned section of the instructional material relevant for answering the exam question.

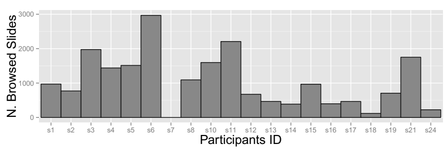
6 Results

6.1 Descriptive Data

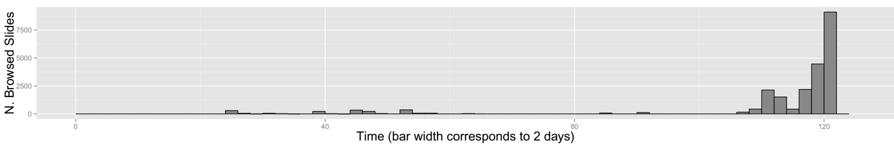
We first report the quantitative data on the classroom participation in producing, sharing and browsing annotations. On average each student annotated 36.05 (SD = 29.77) slides (10% of the total material). Considering the number of browsed annotated slides, each user accessed 1182 (SD = 920) annotated slides. The user-based histograms showing the number of shared annotated slides in Figures 5(a) and the number of browsed slides in 5(b) suggest that almost all users participated in the usage of *annOot*. It is interesting to notice that those students who did not take or share any note still used *annOot* for browsing the slides annotated by the other students. In the time-based histogram in Figure 5(c) we show that the usage of the browsing tool was scarce during the semester and skyrocketed in the two weeks before the exam.



(a) Shared annotated slides per user.



(b) Browsed slides per user.



(c) Time-based histogram of the number browsed annotated slides.

Fig. 5. Descriptive statistics of students participation with *annOot* system

6.2 The Synergistic Effect between Note-Taking and the Note-Browsing

We here address the research hypothesis H_1. Linear mixed effect regression was employed to measure the relationship between the number of annotated slides per lecture ($NumAnn_{LEC}$) and the browsing time of the annotated slides of the same lecture (ABT_{LEC}). While the first activity is done in classroom, the second one is an out-of-classroom learning experience. Given the nature of the test the user identifier was taken into account as grouping factor. The resulting linear model exhibits a significant intercept ($Est. = 688$, $Std.E. = 268$, $t(57) = 2.57, p = 0.013$) meaning that each user on average spends 11 minutes (688 sec) in browsing the annotated slides of one specific lecture. Also the slope resulted do be positive and significant ($Est. = 54$, $Std.E. = 15$, $t(57) = 3.53$, $p = 0.0008$) indicating that the ABT is extended by roughly one additional minute (54 sec) multiplied for the number of shared annotations during that lecture. We accept H_1.

6.3 Student Examination Performance

We here address the research hypotheses H_2 and H_3. For each student we took into account the two questions, $Q = \{q_i, q_j\}$, asked during the exam. We used the bipartite graph $G_Q = (Q, L, E)$ to select the subset of the instructional material $I\{i_s, i_t, ..., i_u\}$, which understanding was required to correctly answer to those two questions (as already explained in Section 5.3). Then for each student, we recalculated the number of annotated slides and the annotated slide browsing time relative to such partition of the instructional material, and we labeled the new variables respectively $NumAnn_{REL}$ and ABT_{REL}.

A Kendall rank correlation test between $NumAnn_{REL}$ and the exam score was carried out[2]. The result is not significant ($\tau(18) = -0.10$, $p = 0.53$). We reject H_2. A second correlation test between ABT_{REL} and the exam score was conducted. For this second test we found a positive and significant correlation ($\tau(18) = 0.32$, $p = 0.042$). We accept H_3. This result encouraged us to elaborate a deeper investigation. We questioned whether the time spent browsing only the annotations produced by the best students of the classroom better correlates with the individual exam score. Therefore the student group was split in two sub-groups using the median of the exam score ($Q2_{ExamScore} = 4.95$). For each student, the browsing annotation time was recomputed considering only the annotations produced by the best half of the classroom; the new variable is labeled as $ABT_{REL,BEST}$. A third correlation test between $ABT_{REL,BEST}$ and the exam score was conducted. The test revealed a strong positive and significant correlation ($\tau(18) = 0.35$, $p = 0.032$). We can conclude that browsing the annotations produced by the best-in-classroom students is correlated with the individual exam grade.

[2] Kendall rank correlation test is a non-parametric test that does not assume any assumption related to the distributions and is robust to the presence of outliers. We opted for this correlation method because the Shapiro and Wilk normality test revealed that the variable exam score is not normally distributed ($W = 0.88$, $p = 0.019$) and that the low sample size of the dataset ($N = 20$) could exhibit a higher exposure to the presence of outliers.

6.4 Friendship Bias in Browsing Shared Annotations

When the user enters a new slide he or she can access to someone's annotations (produced and shared by *ego*, *friends* or *non-friends*, see Section 5.3) by clicking on her name. As stated in H_4, H_5 and H_6, we expect a social bias conveyed through the name list of annotators (see panel 5 in Figure 2). This bias can influence the user's decision of whose annotations should be accessed first. To verify this we compared the theoretical probabilities of randomly first clicking on one's own annotations (*ego*), on friends' annotations (*friends*) and on non-friends' annotations (*non-friends*) with the related measured empirical probabilities. For instance we computed the theoretical and empirical probabilities of first clicking on a friend's annotation on a specific slide as follows:

$$P(friends) = \frac{|friends|}{|annotations|} \qquad F(friends) = \frac{n_{friends}}{n_{annotations}}$$

Considered a specific annotated slide, the theoretical probability of first clicking on a friend's annotation $P(friends)$ corresponds to the number of annotations produced by the user's friends ($|friends|$) divided by the total number of annotations produced by everybody for that slide ($|annotations|$). The empirical probability of first clicking on a friend's annotation $F(friends)$ corresponds to the number of observed first clicked friends' annotations in that slide ($n_{friends}$), divided by the total number of first clicked annotations ($n_{annotations}$). Only the data entries in which the user could click on all the mentioned social categories were considered for this test. To measure any discrepancy between the empirical and theoretical probabilities, we employed linear mixed effect regression with probabilities as measured variable and with the categorical predictor indicating whether they are theoretical or empirical. Three models were built, one for each type of social category. A nested grouping factor accounting the slide identifier was introduced in the models. The empirical probability of first clicking on one's own annotations is higher ($F(Ego) - P(Ego) = 0.11$) than the correspondent theoretical probability ($t(433) = 5.69$, $p < 0.0001$). H_4 is accepted. The empirical probability of first clicking on one's friends annotations is not significantly different ($F(friends) - P(friends) = 0.0049$) than the correspondent theoretical probability ($t(433) = 0.23$, $p = 0.81$). H_5 is rejected. The empirical probability of first clicking on *non-friends*' annotations is lower ($F(Others) - P(Others) = -0.12$) than the correspondent theoretical probability ($t(433) = -5.43, p < 0.0001$)). H_6 is accepted.

As last step of the analysis we tested the research hypotheses H_7 and H_8. We investigated whether the social provenience of an annotation can influence its browsing time. Therefore we constructed a linear mixed-effects model where the *ASP* categorical variable predicts the *ABT* numerical variable. The user identifier was taken into account as a grouping factor. The results show that the user spends less time on her/his own annotated slides (ABT_{ego} - $ABT_{friends}$ = -2.4 sec) than on friends' ones ($t(11884) = -2.34$, $p = 0.019$). The user spends less time on non-friends' annotated slides ($ABT_{non-friends}$ - $ABT_{friends}$ = -1.7 sec) than on friends' ones ($t(11884) = -2.71, p = 0.0067$). We accept H_7. However the time spent non-friends' annotated slides is not different ($ABT_{non-friends}$ - ABT_{ego} = 0.82 sec) than the time spent on one's owns annotated slides ($t(11884) = -0.88, p = 0.37$). We reject H_8.

7 Discussion and Conclusion

From the quantitative observations on the usage of the tool, we can conclude that students used the tablet for note-taking fairly regularly during the arc of the semester. This strengthen existing experimental and observatory investigations on the ability of tablets at mimicking the physical properties of paper in digital annotations [33,14]. We also observed that students participated into the note sharing process and in browsing the shared annotations. Therefore our tool implemented Marshall's idea of mobile technology able to mediate the transition *From Personal to Shared Annotations* [20]. In addition we observed that *annOot* spontaneously broke the barrier of the in-classroom learning activities, since it was used by the students as complementary resource for the exam preparation. We also measured a synergistic effect between sharing the notes taken during a lecture and the time spent browsing the slides of the same lecture. We gave two plausible interpretations to this phenomenon: 1) taking and sharing notes increases students' engagement and motivation in browsing the notes taken from other classmates. 2) highly engaged students tend to browse often the course material and consequently take and share more notes. We are not able to discriminate any of these 2 interpretations.

We did not measure any effect of taking notes in classroom, using a tablet, on the exam grade. This result is interesting if considered that a vast research literature showed the positive effects of self-note-taking on learning performance [30,5,12]. Clearly the limitations of the study prevent us from taking any strong conclusion on this result. However one interpretation could be that there is still room for improvement for tablet-based note-taking technologies. One immediate solution, easily integrable with our technological architecture, would rely on the integration of paper-based interfaces in *annOot* system [32,27,13]. For instance the students could take notes on printed handouts and decide to share them by taking photos of the annotated areas using any mobile device. The recognition system integrated in *annOot* [23], can already robustly deal arbitrary perspectives in the document image.

Further on we showed how the time spent browsing the shared annotated instructional material before the exam, is positively correlated with the final exam score. This result is in line with a vast literature focusing on the effects of external and multiple representations on leaner's recall performance [31,17]. We would like to complement this result with an additional interpretation: the awareness of other students' annotative productions can disclose the emergence of a learning disequilibrium due to socio-cognitive interactions [26]. For instance the awareness of friends' and non-friends' annotations, can induce the students to improve, adapt, or complement their annotations and annotating behaviours. One interesting feature that should also be taken into account is the instructors' feedback. This type of feedback could reveal which passages of the instructional material are not well understood by the students. As already experienced by Crossen et al. [6], such feedback could result in modifications in the teaching work flow. Using our system the teacher can also easily integrate relevant notes to the teaching material of the subsequent year, as suggested by David et al. [8]. In addition we refined this latter result by observing that the time spent browsing the slides annotated by the best students of the classroom better correlates with the exam grade. Such an expertise bias could drive prioritization criteria in the visualization of the annotations.

As last result we observed how the social provenience of shared annotations can affect the user browsing behaviour. The user tends to first click more often on his or her own annotations and less often on non-friends' annotations. We also found that the user spends more time on friends annotations than his owns' or the others'. This result shows that the students' browsing of socially labeled notes is biased in favour of strong ties; this means that *annOot* system modulates the interaction between the student and the annotated material concordantly with the classroom social boundaries. This modest manifestation of the *homophily principle* in social network analysis [21] could be compensated, in the future, by using content-based recommendation systems.

We conclude that the discussed results will encourage the deployment of *annOot* system as fully accessible University micro-publishing platform based on both digital and paper note-taking. We believe that the system can support fruitful learning activities, complementary to traditional means of lecturing. An evolution of the system should integrate a social network of the students where each attended course generates a potential circle of friends. In addition it should integrate a recommender system of authored or annotated course material, based on contents and students' proficiency.

Acknowledgements. David Bréchet has our utmost gratitude for his insightful assistance pertaining to information technology. We also thank all the students who participated in the study. This work is funded by NCCR IM2 (Interactive Multimodal Information Management, IM2-IP2).

References

1. Anderson, R., Davis, P., Linnell, N., Prince, C., Razmo, V., Videon, F.: Classroom presenter: Enhancing interactive education with digital ink. Computer 40, 56–56 (2007)
2. Berque, D.: An evaluation of a broad deployment of DyKnow software to support note taking and interaction using pen-based computers. Journal of Computing Sciences in Colleges 21, 204–221 (2006)
3. Bonnard, Q., Verma, H., Kaplan, F., Dillenbourg, P.: Paper Interfaces for Learning Geometry. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 37–50. Springer, Heidelberg (2012)
4. Chi, M.T.H.: Quantifying qualitative analyses of verbal data: A practical guide. Journal of the Learning Sciences 6, 271–315 (1997)
5. Craik, F.I.M., Lockhart, R.S.: Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior 11, 671–684 (1972)
6. Crossen, A., Budzik, J., Warner, M., Birnbaum, L., Hammond, K.J.: XLibris: an automated library research assistant. In: Proceedings of IUI 2000, pp. 49–52. ACM Press, New York (2001)
7. Davis, J.R., Huttenlocher, D.E.: Shared annotation for cooperative learning. In: Proceedings of CSCL 1995, pp. 84–88. L. Erlbaum Associates Inc., Hillsdale (1995)
8. Davis, R.C., Landay, J.A., Chen, V., Huang, J., Lee, R.B., Li, F.C., Lin, J., Morrey III, C.B., Schleimer, B., Price, M.N., Schilit, B.N.: NotePals: Lightweight note sharing by the group, for the group. In: Proceedings of the CHI 1999, pp. 338–345. ACM Press, New York (1999)
9. Fischler, M.A., Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Comm. ACM. 6, 381–395 (1981)

10. Fu, X., Ciszek, T., Marchionini, G., Solomon, P.: Annotating the Web: An exploratory study of Web users' needs for personal annotation tools. In: Proceedings of ASIST 2005 (2005)
11. Guimbretire, F.: Paper augmented digital documents. In: Proceedings of UIST 2003, pp. 51–60. ACM Press, New York (2003)
12. Hartley, J., Davies, I.: Note-taking: A critical review. Innovations in Education and Teaching International 15, 207–224 (1978)
13. Heiner, J.M., Hudson, S.E., Tanaka, K.: Linking and messaging from real paper in the Paper PDA. In: Proceedings of UIST 1999, pp. 179–186. ACM Press, New York (1999)
14. Hu, W.: More Schools Embrace the iPad as a Learning Tool. The New York Times (2011)
15. Kam, M., Wang, J., Iles, A., Tse, E., Chiu, J., Glaser, D., Tarshish, O., Canny, J.: Livenotes: A system for cooperative and augmented note-taking in lectures. In: Proceedings of CHI 2005, pp. 531–540. ACM Press, New York (2005)
16. Krackhardt, D.: Cognitive social structures. Social Networks 9, 109–134 (1987)
17. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. Cognitive Science 11, 65–100 (1987)
18. Liao, C., Guimbretire, F., Hinckley, K., Hollan, J.: Papiercraft: A gesture-based command system for interactive paper. ACM Transactions on Computer-Human Interaction 14, 18:1–18:27 (2008)
19. Marshall, C.C., Price, M.N., Golovchinsky, G., Schilit, B.N.: Introducing a digital library reading appliance into a reading group. In: Proceedings of DL 1999, pp. 77–84. ACM Press, New York (1999)
20. Marshall, C.C., Brush, A.J.B.: From personal to shared annotations. In: CHI 2002 Extended Abstracts, pp. 812–813. ACM Press, New York (2002)
21. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology 27, 415–444 (2001)
22. Miura, M., Kunifuji, S., Shizuki, B., Tanaka, J.: AirTransNote: augmented classrooms with digital pen devices and RFID tags. In: Proceedings of WMTE 2005, pp. 56–58. IEEE Computer Society, Washington (2005)
23. Nakai, T., Kise, K., Iwamura, M.: Use of affine invariants in locally likely arrangement hashing for camera-based document image retrieval. In: Bunke, H., Spitz, A.L. (eds.) DAS 2006. LNCS, vol. 3872, pp. 541–552. Springer, Heidelberg (2006)
24. O'Hara, K., Sellen, A.: A comparison of reading paper and on-line documents. In: Proceedings of CHI 1997, pp. 335–342. ACM Press, New York (1997)
25. Ovsiannikov, I.A., Arbib, M.A., McNeill, T.H.: Annotation technology. International Journal of Human-Computer Studies 50, 329–362 (1999)
26. Piaget, J.: The psychology of intelligence. Routledge, London (2004)
27. Rao, R., Card, K., Johnson, W., Klotz, L., Trigg, R.: Protofoil: storing and finding the information worker's paper documents in an electronic file cabinet. In: Proceedings of CHI 1994, pp. 180–185. ACM Press, New York (1994)
28. Schilit, B.N., Golovchinsky, G., Price, M.N.: Beyond paper: Supporting active reading with free form digital ink annotations. In: Proceedings of CHI 1998, pp. 249–256. ACM Press, New York (1998)
29. Sereno, S.C., Rayner, K.: The when and where of reading in the brain. Brain and Cognition 42, 78–81 (2000)
30. Di Vesta, F., Gray, G.S.: Listening and note taking. Journal of Educational Psychology 63, 8–14 (1972)
31. Waller, R.: Understanding network diagrams, Los Angeles (1981)
32. Wellner, P.: Interacting with paper on the DigitalDesk. Communications of the ACM 36, 87–96 (1993)
33. Wolfe, J.: Annotation technologies: A software and research review. Computers and Composition 19, 471–497 (2002)

Interdisciplinary Cohesion of TEL – An Account of Multiple Perspectives

Philip Meyer[1], Sebastian Kelle[2], Thomas Daniel Ullmann[3], Peter Scott[3], and Fridolin Wild[3]

[1] University of Augsburg, Media Lab, Augsburg, Germany
philip.meyer@phil.uni-augsburg.de
[2] Stuttgart Media University, Stuttgart, Germany
kelle@hdm-stuttgart.de
[3] The Open University, Knowledge Media Institute, Milton Keynes, UK
{t.ullmann,p.scott,fridolin.wild}@open.ac.uk

Abstract. Research areas and academic disciplines are not static: they change over time with new strands emerging and old ones disappearing. Technology-enhanced learning is a relatively young field of academic activity, getting more broad in scope as it matures. In this paper we seek to assess the state of interdisciplinarity in this academic community, presenting the findings of a quantitative study on mutual engagement, shared practices and methodologies, and sense of joint enterprise via a European research network in between learning and technology disciplines. An exploratory cluster analysis is used to identify different stakeholder groups in technology-enhanced learning research and a social network analysis shows how these are connected to each other. Statistical analysis suggests that a multidisciplinary workplace and study background of researchers are major influencing factors for the choice of border-crossing methodology and terminology. Additionally, results from a supplementary survey on the interdisciplinary cohesion between the fields of technology-enhanced learning and educational development support the view that pedagogical and technological sub-disciplines highly intersect in this field.

Keywords: Interdisciplinarity, Research, Methodology, Technology-Enhanced Learning, Educational Development, Social Network Analysis, Science Studies.

1 Introduction

What constitutes an academic discipline is itself subject to scientific debate [1]. There are, however, elements that often are connected to it – such as the presence of a community of scholars, a tradition in theory and methodology, agreement on what constitutes new knowledge, or the existence of a communication network [1]. In this work, we investigate the nature of interdisciplinarity of the field of technology-enhanced learning (TEL), as well as the cohesion between TEL and Educational Development (ED) communities. We utilise a survey among the members of the community platform of the European Network of Excellence STELLAR [2] for this

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 219–232, 2013.

and backup findings about community structure with a social network analysis of the professional relationships of its members. In a final step, we used an additional survey to gain insight into mutual interconnection and awareness between TEL and ED perspectives.

In scientific theory, interdisciplinary approaches to problems are often discussed as providing most appropriate solutions [3-4]. It is only the combined expertise of multiple disciplines, which can help to develop solutions that address all relevant aspects of a problem and provide an overarching perspective. The question remains how education and learning can be best facilitated in a world where communication, as a means for learning in a constructivist fashion, is increasingly technology-mediated. Still, social science oriented educational researchers and practitioners do not necessarily consider themselves members of the TEL field, but may prefer to locate their work in the area of educational development. This field has evolved quite independently of technological innovations and has only recently confronted itself with the multitude of new learning possibilities driven by the TEL community.

To approach these aspects, the rest of the article is structured as follows: First, we outline the study design to then present results on research traditions and academic knowledge production. A cluster analysis allowed for describing methodologically distinct parts of the TEL community. Those parts are connected to each other through interpersonal relationships, which have been identified by a social network analysis. Subsequently, the results are discussed and conclusions for the further development of the field of technology-enhanced learning are derived. Finally, we present results from a supplementary survey on mutual awareness and reception of the stakeholders of TEL and ED. The participants of this survey were drawn from the International Consortium of Educational Development (ICED) as well as from the European Association of Technology Enhanced Learning (EATEL).

2 Study Design

Traditionally, disciplines are measured via collaboration indicators, including social media relationships, co-authorship and co-citation networks, particularly when investigating interdisciplinarity [5-6]. In our work, we have instead designed an actor-centred survey to investigate individual backgrounds, practices, and perceptions of researchers in the field. The analysis of cross-disciplinarity is complemented with a social network analysis of the connections set within the open community platform of the European Network of Excellence STELLAR[1], where researchers connect to each other using the friendship functionality. The general question investigated in this paper addresses the nature and strength of a TEL community of practice [7] with mutual engagement, shared practices, and a sense of joint enterprise.

Three analytical instruments have been combined in this study. First is the analysis of group differences with reference to the individual disciplinary study background. For this, data was collected with a survey sent via e-mail in June 2011. The poll was open for three weeks, and from 1,200 contacted persons, 123 completed the

[1] `www.teleurope.eu` – the "help" section provides an extensive overview of the community platform's functionalities.

questionnaire. The sample is representative for the European TEL network platform with its approx. 1,200 members at the time of study. The three broad groups of analysis are those with a 'social science' background, including all social science and humanities disciplines[2], those with an 'engineering' background, including (and to a great extent consisting of) computer science, and those with 'multidisciplinary background', containing researchers who have been trained in both disciplinary fields. Participants were asked to position themselves in these broad study fields. Such a study conception comes with the advantage of having distinct comparable groups, hence blinding out the variety of interdisciplinary fields, which represent TEL. Starting from these 'primeval' academic fields of expertise, the approach at hand makes it possible to identify alternative viewpoints, deconstructing the perceived dualism between technology and learning – engineering and social sciences.

As a second step, the sample was re-grouped via those variables that deal with knowledge-related practices, scientific attitudes, interdisciplinary identification, and use of theories, methods, publication practices, etc. This is achieved through the method of hierarchical cluster analysis [9], resulting in clusters of researchers that share similar characteristics. The clusters allow for describing the variety of interdisciplinary subgroups, thus raising awareness of structures in the community that remained concealed before applying the cluster analysis.

In a third step, the focus was set on 'integration'. This meant looking at how persons in the different, meaningful clusters are connected on the Internet. Here, the method of choice is social network analysis (SNA) [10], which allows looking at intra- and inter-cluster relations. The SNA builds upon the results of the previous analyses and completes the process of deconstructing disciplinarity.

The complementary study on cohesion between TEL and Educational Development is described in more detail in Section 6.

3 Research Practices and Theoretical Stances

3.1 Technology-Enhanced Learning Traditions

The observed sample has four main backgrounds: The social scientists, including also humanities, the engineering scientists, the ones who did study in both fields, and consequently the ones who come from other areas such as natural or life sciences. Social science researchers make up for almost half of the community (43%), engineering scientists one third. The rest are 'multidisciplinarians'[3], which is defined through an academic background related to both social science and engineering (26%). Professionals, who took a new career path in TEL research and come from disciplines that are not genuinely social science and technology related are rare.

All four backgrounds are equally represented through senior researchers holding a professorship. It is noteworthy that having a multidisciplinary background correlates with a higher age. Researchers aged 30 years and younger are rarely trained in multiple disciplines (12%). Looking at the nature of institutions where TEL research

[2] Following the subject area classification published by the German Research Foundation [8]

[3] The naming does not conclude that they perceive themselves as multidisciplinarians. It is solely derived from the question "In which scientific fields have you been studying?"

is being conducted, most of them are linked to a single discipline. Still, one in four researchers works in a multidisciplinary unit. Both computer and social scientists work together mainly with colleagues from the same discipline (65%; 72%).

It is not necessarily the case that persons, who studied more subjects, also perceive their studies as more interdisciplinary. On the contrary, a majority of researchers trained in multiple disciplines feel that their studies were rather made of "several unconnected programmes" (58%). Across all disciplines a strong commitment to interdisciplinarity as a "mode 2 of work" [11] is visible, 78% even prefer it to working in one single discipline. "Mode 2", which started to emerge from the 1950s on, is characterised by problem-focused, context-driven, and interdisciplinary work. It involves teams, which are brought together for short periods of time to work on specific problems in the real world. From our findings, researchers trained in multiple disciplines tend to be more confident in saying that they "bridge different scientific fields" (e.g. Humanities & Engineering Sciences) in their research practice and use of methods (Mann-Whitney $U = 1032$, $p = .014$).

Technology-enhanced learning is rooted in empirical methodology. Table 1 shows that almost every researcher uses qualitative and quantitative empirical methods for research. Also common across the disciplines is design methodology, namely user-centred design and design-based research, as well as socio-cognitive engineering. Other methods are only practiced by some disciplines: Computer scientists prefer theoretical and experimental computing methods, while social scientists have a propensity for ethnographic methodology. For the latter, a significant share of researchers had not even heard of it (15% of the total saying "I don't know this method").

Technology-enhanced learning as a research field is highly influenced by the "community of practice" concept [7]. It is a theoretical perspective, along which most of the European TEL researchers across all disciplines base their inquiries (see Tab. 1, left). As the theory is related to the tradition of socio-cultural constructivist learning, not surprisingly many researchers also favour this theoretical approach. Other learning theories like cognitivism and behaviourism are less common and mostly referred to by social scientists.

Table 1. Left: "On which of the following theoretical perspectives do you base your research?" Right: "Which of the following methods do you use in your research?"[4]

Theory	p-Value	Agreement*
Computer Science Theories: Theory of computation	<.01	33%
Computer Science Theories: Representation theory	*ns*	33%
Computer Science Theories: Artificial intelligence / MLT	<.001	50%
TEL Theories: Communities of practice (CoP)	*ns*	88%
TEL Theories: Actor-network theory (ANT)	*ns*	51%
TEL Theories: Constructionism	*ns*	68%
Learning Theories: Constructivism	$.06^{ns}$	87%
Learning Theories: Cognitivism	< .001	77%
Learning Theories: Behaviourism	.04	55%

* At least "Seldom"

Methodology	p-Value	Agreement*
Theoretical computing methods	.03	43%
Modelling and simulation methods	*ns*	67%
Experimental computing methods	.01	58%
Design (based) research methods	*ns*	83%
User-centered design methods	*ns*	87%
Socio-cognitive engineering methods	*ns*	64%
Quantitative empirical methods	< .01	93%
Qualitative empirical methods	< .001	93%
Ethnographic methods	< .01	37%

* At least "Seldom"

[4] Brief method / theory descriptions were given to the survey participants, in order to avoid misunderstandings.

The second most cited interdisciplinary theory after ‘community of practice’ is ‘constructionism’[12], followed by ‘actor-network theory’ [13], which is listed by around half of all researchers. Also, all groups refer to ‘representation theory’. Only a minority of TEL researchers base their research on computing theories. 50% are somehow familiar with artificial intelligence and machine learning theory, while theories of computation and representation are each only taken into account by one third of the researchers – significantly more by computer science researchers.

3.2 Publishing Culture: Academic Knowledge Production

TEL Researchers cite the differing language and terminology as the main barrier to interdisciplinarity [14]. In this survey, researchers were asked what they understand by certain ambiguous terms like ‘scenario’, ‘intervention’, and ‘evaluation’, (being elementary concepts in the field of technology-enhanced learning). Contrasting understandings were evident, in a sense that the study background correlates with the choice of terminology. TEL researchers, who studied social science, state that they use the term ‘evaluation’ for the evaluation of people, like teachers, students, employees, and others. Researchers with a computer science background in contrast often evaluate technology and “the performance of hardware and software”. A Kruskal-Wallis test revealed significant differences $H(2, N=116) = 9.81$, $p < .01$. Many multi-disciplinarians chose a system-oriented meaning of the term, evaluating “a system involving people and technology” (H=4.33, *ns*). Similar results occurred for other terms. Persons preferred different definitions, depending on their disciplinary background. In general, most researchers chose system-oriented term meanings, they account for half of all answers (50%).

As Open Access to publications can help to establish more interdisciplinarity, the study at hand was also looking at publishing of research results. Therefore, results of the DFG Open Access study from 2008, which was conducted for the German academic landscape [8], could to be reproduced on a European Scale (see Tab. 2). Aim of the study was to see how widespread different forms of Open Access are taken up across the disciplines and which formats are favoured. Concerning the internationality of research, three quarters of the community address researchers outside their work country at least “usually”. Differences between the groups can be found, as people with a social science background seem to target international

Table 2. Comparison with results of the DFG Open Access study [8]

Publication audience	DFG Open Access study (2005)			TEL interdisciplinarity survey		
	Social science	Engineer. science	Total	Social science	Engineer. science	Total
Researchers from own discipline	99%*	99%*	99%*	88%*^{3}	97%	91%^{84}
Researchers from other disciplines	55%*	41%*	48%*	43%	46%	45%
Application-oriented target groups	14%*	**39%***	18%*	37%	36%	38%
Interested non-professionals	**16%***	1%*	5%*	**37%**	15%	29%
International audience	43%*	**77 %***	78%*	57%*^{2}	**92%***^{2}	76%*^{2}
* at least “predominantly” *^{2} at least “usually” *^{3} no significant difference	N=235	N=225	N=1.023	N=51	N=39	N=123

audiences less often (Mann-Whitney U = 1209, p < .01). Almost everyone in the engineering group has a strong international focus, as opposed to only around half of the social science group. Similar to the DFG study, social scientists target non-professionals more (U = 1305, p = .02). Audiences explicitly named here are, e.g., teachers, particularly secondary school and university teachers, students, company trainers, policy makers, and academic administration. Different from the original study, social scientists in technology-enhanced learning have a more application-oriented focus than the 'conventional social scientists' of the DFG study. They target practitioners as much as engineering researchers do.

4 The Parts of the Community: Cluster Analysis

4.1 Method of Clustering

The aim of the cluster analysis was to look at the domain of technology-enhanced learning without the dichotomy of two large disciplinary fields. Figure 1 shows the clusters, which resulted from the hierarchical cluster analysis, using a between-group comparison algorithm along the rescaled Pearson correlation. A cluster was accepted, if it consisted of at least N=5 cases with a maximum rescaled distance of 14. This led to eight clusters A-G (N=101), with 22 cases being neglected, as they did not cluster sufficiently well. This cluster analysis allows for a better differentiation between stakeholder groups in TEL, looking at all surveyed variables and not only the study background. Two new aggregate variables were calculated in order to visualise the clusters, namely *V25 TEL participation* and *V26 disciplinary orientation*.

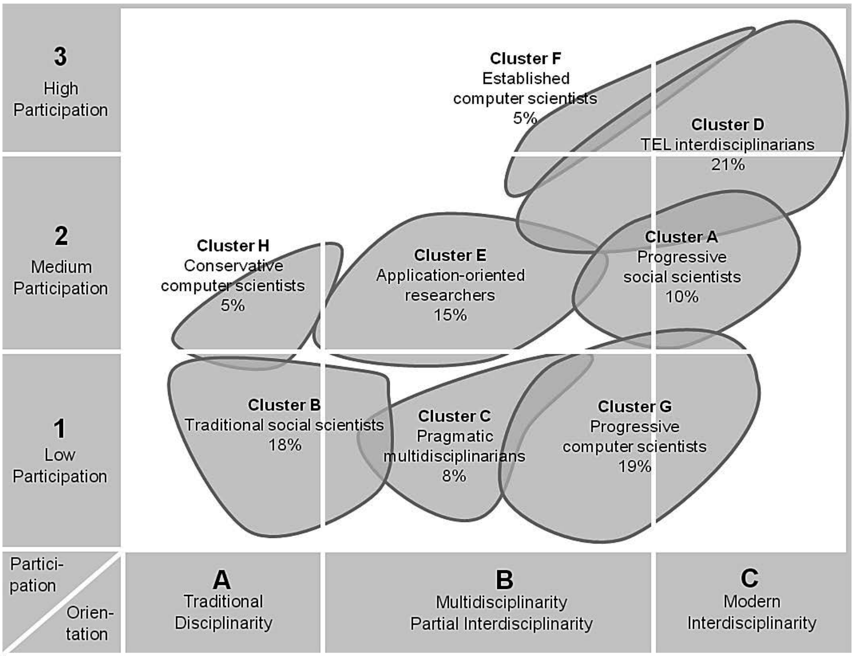


Fig. 1. Clusters of the European TEL community (for detail see section 4.2)

'TEL Participation' includes all variables, which indicate that a researcher is actively participating in TEL research. It sums up the score of activities, the involvement in core research areas, the general use of and knowledge about methods, as well as theories, the targeting of multiple audiences, the publication score, project participation, and platform connections.

The dimension 'Disciplinary orientation' takes into account all variables, which are related to the concept of a strong, integrative interdisciplinarity. This includes having a study background in several disciplines, perceiving one's own studies and work as very interdisciplinary, having a positive attitude towards interdisciplinarity, using possibly discipline-bridging terminology, methodology and theory, targeting audiences other than one's own, publishing open access, and working together with multidisciplinary colleagues. The clusters are named in this figure based on an analysis of specific cluster traits.

4.2 The Stakeholders of Technology-Enhanced Learning Research

The cluster analysis revealed several groups of TEL stakeholders, which differ from each other. Differences are sometimes marginal and should only be interpreted as indicators for further, more representative research. Still, the statistics in table 3 hold potential for reflection on the interdisciplinary nature of European TEL research.

A striking gap, in the figure 1 map top left corner, contains no clusters. It indicates that there are very few traditional disciplinary researchers, who highly participate in European TEL. For the social sciences this may relate to fewer international bonds [15]. In a Europe-wide study on mobility and internationality of researchers it turned out that social scientists focus strongly on "national disciplinary traditions and histories", finding it more important to work together with their disciplinary colleagues than in an interdisciplinary way. Looking at traditional computer scientists, one explanation might be the perceived lack of respect towards the *science* part of computer science within the TEL community [14], treating it as a functional "service element", not noticing computer scientists' need for clear definitions [14]. This could explain less participation from the aforementioned groups.

Table 3. Cluster statistics

	A: 'Progressive Social Scientists' (10%)	B: 'Traditional Social Scientists' (18%)
Location	university-based, not necessarily in social science institutions	social science institutions at universities
Sex	predominantly female	both
Profession	researchers and post-docs	young researchers and PhDs
Age	30 to 40	< 30
Background	only social science	only social science
Activities	coaching, teaching, policymaking, ICT work	no programming, teaching only rarely
Network	loosely connected in the European network	well connected, but not to cluster A →

→ Terminology	mostly systemic (involving people and technology)	people-oriented (less focus on technology)
Methodology	theoretical, experimental computing, and modeling methods, empirical methods, no ethnography	social science methodology, quant. & qual. empirical methods, design research
Theories	connectivism, cognitive learning, computing theory	cognitivist and constructivist learning theo. no representation theory
Publishing	national academic audiences	academic audiences; publish less in journals
TEL areas	many, not involved in workplace learning	few, formal & informal learning, CSCL
Attitude	mainly identify with social science; study background is perceived as inter-disciplinary (ID); interdisciplinary research is something 'normal'; prefer ID research from working monodisciplinary	do not perceive their study background or their current work as very interdisciplinary; interdisciplinary research seems hard to achieve and publish
	C: 'Pragmatic Multidisciplinarians' (8%)	**D: 'TEL Interdisciplinarians' (21%)**
Location	social science institutions at universities	interdisciplinary university institutions
Sex	both	both
Profession	experienced researchers	research associates and lecturers
Age	40 to 50	30 to 40
Background	social science as well as computer science	social science as well as computer science
Activities	programming above average	policymaking & programming above avg.
Network	connect a lot of clusters	well connected
Terminology	mostly systemic	mostly systemic
Methodology	empirical methodology, more qualitative, design-based research	trained in computer and social science; use design-based research, user-centered design, some socio-technical engineering
Theories	often communities of practice	theories of computing and learning
Publishing	international academic audiences; conventional, few open formats	also applied audiences, non-professionals; a lot in conventional and open formats
TEL areas	few, but mostly formal and mobile learning	many, formal/informal learning; workplace learning; personalised, contextualised l.
Attitude	mainly identify with social science; background is perceived as inter-disciplinary; think ID research is hard to achieve	see themselves as 'interdisciplinarians'; state that TEL is interdisciplinary; "interdisciplinarity is something normal"
	E: 'Application-oriented researchers' (15%)	**F: 'Established Computer Scientists' (5%)**
Location	school- or company-based	university-based
Sex	both	predominantly male
Profession	experienced teachers and researchers	professors and research associates
Age	40 to 50	40 to 50
Background	largely social scientific background	computer science
Activities	policymaking, practice less research than others	policymaking more than average
Network	less central in the TELeurope network	highly active, involved in many EU-Projects
Terminology	people-oriented (less focus on technology)	choose more people-related terms
Methodology	no relation to computer science, use modeling methods, qualitative research	methods from computer and social science; prefer qualitative methodology to quantitative; ethnography; some socio-technical engineering
Theories	mostly constructivism	familiar with most kinds of theories
Publishing	practitioners and non-professionals as audience; publish mostly in open formats	publish more than others; very international target other disciplines
TEL areas	workplace learning; formal/informal learning; digital divide	CSCL, formal learning and interoperability
Attitude	mainly identify with social science; Interdisciplinarity is normal to them, they don't think it is hard to achieve/publish	prefer working interdisciplinary from working in a single discipline; don't identify much with their computer science background
	G: 'Progressive Computer Scientists' (19%)	**H: 'Conservative Computer Scientists' (5%)**
Location	university-based	university-based
Sex	both	mostly female
Profession	research assistants and PhDs	research assistants and PhDs
Age	< 30	< 30
Background	computer science	computer science
Activities	program. above avg.; no coaching, no policy-making	teaching and ICT work above average
Network	most connected cluster in the European TEL network	relatively unconnected, few European projects
Terminology	systemic, discipline-bridging terminology	technology-oriented terminology
Methodology	compared to other computer scientists they use more (quantitative) empirical methods	mainly computing methods and some user-centered design
Theories	computer science; not very used to learning theories	also interested in learning theories constructionism, cognitive/machine learning →

→ Publishing	target application-oriented audiences; publish seldom in journals and open formats,	only address academic audiences; active with publishing in international journals
TEL areas	few, mostly personalisation and interoperability	especially formal learning, interoperability and personalization
Attitude	For them studying computer science meant studying an 'interdiscipline'; They consider themselves as interdisciplinarians	don't see themselves as 'interdisciplinarians'; state that TEL is not necessarily interdisciplinary; don't always see ID as something useful

5 The Communication Network: Social Network Analysis (SNA)

On the TELeurope.eu platform, it is possible to add other members as friends. No confirmation is necessary, even 'one-way' connections count as friendships. For our investigation, a network analysis was conducted, in order to see how researchers from different disciplines are connected to each other in mutual engagement. Two thirds of the participating researchers on TELeurope.eu are connected to others. The majority thereof has no more than 20 connections. Engineering science researchers are most connected (87%, Mann-Whitney U= 1140, $p < .01$), followed by multidisciplinary researchers (68%). Social scientists are least connected (57%). There are also differences along age, as younger researchers are over-represented[5], and location, such as e.g. French researchers being under-represented.

The degree of connectedness is very different for the identified clusters. The 'progressive computer scientists' of cluster G are most connected, with an average degree level *dc* of 3.3 (see Fig. 3). High values are present for any measure of centrality. Even though their general research participation is lower than in other clusters, the contributing researchers are active networkers and connect with all clusters. Another computer science cluster, the 'established computer scientists' (F) is the second most connected one, despite the relatively high age. They are mainly linked to the aforementioned progressive computer science researchers (see Fig. 2).

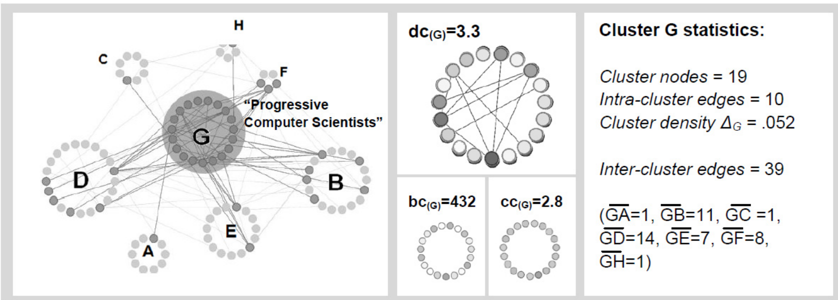


Fig. 2. Example of a Cluster Network Reporting from the study

[5] This may, however, reflect the typical pyramid structure of institutional organization: few, typically older upper level staff, more, typically younger mid- and lower level.

The strongly participatory ‘TEL interdisciplinarians’ (D) stand out, as they have the highest betweenness centrality value. This indicated that many others are dependent on this group, in order to reach indirect contacts. Even though it does not have a lot of direct (*dc*=1.6) and indirect (*cc*=2.2) connections, cluster D is connected to all other clusters. TEL interdisciplinarians are not very connected within their cluster. Instead, lots of connections are shared with the ‘applied researchers’ (n=5), the ‘progressive computer scientists’ (n=6) and the young, ‘traditional social scientists’ (n=5).

‘Traditional social scientists’ (B), as identified by the cluster analysis, are neither very interdisciplinary oriented, nor actively participating in TEL research. Surprisingly, they are relatively well connected on TELeurope. They share only a few internal connections and less external connections with multidisciplinarians than others, but are often interacting with the computing researchers in cluster G. The ‘application-oriented researchers’ (E) are less central in the TELeurope network. Only a few nodes are well connected. They are in particular friends with researchers from the bigger clusters, i.e. the ‘progressive computer scientists’ (n=7), the ‘traditional social scientists’ and the ‘interdisciplinarians’ (both n=5). Persons with a very interdisciplinary attitude and a high participation in TEL research belong to the ‘progressive social scientists’ in cluster A. Their activity in interdisciplinary research doesn’t reflect on the TELeurope platform, as they are rather isolated. Only one of the contributing 10 nodes is connected to other cluster. It lies at hand for this cluster that the social scientists play a less active role in the European TEL network.

6 The External View: TEL and Educational Development

There is a phenomenon in system theory concerning the existence of two different systems that can be described with the same concepts [16]. The field of TEL, however, seems to have the opposite property. While several disciplines evolved into one large ‘system’ where different perspectives and application contexts benefit from each other, another type of ‘system’ exists that works on similar goals, while not necessarily identifying itself as part of the TEL field. Such perception is found in the community of Educational Development (ED), which is a pedagogy-oriented field of practitioners focusing on the enhancement of teaching and learning practices in general. Technology is not the main focus here, because it is seen rather as an auxiliary, albeit with increasing importance. The two fields are represented by two distinct communities that operate internationally, both involved in policy creation for education, just to name one of many things they have in common. Both aspire to be regarded as research disciplines, with a number of research journals and high profile conferences available alike.

Taking into account that it is a matter of controversy, we will refer here to TEL and ED as own fields and/or disciplines. The questions resulting thereof are obvious. Do the two disciplines have enough mutual awareness and collaboration incentive so that they can benefit from each other? What about the perceived separation between the disciplines and their trends of growing together or drifting apart?

To find out more about this, we ran an additional survey targeting members of European Association for Technology Enhanced Learning (EATEL) as well as

International Consortium for Educational Development (ICED). The survey asked the same questions to members of either field.

We asked the participants of the study to choose between these different roles: Technology Enhanced Learning / more technologically oriented discipline; and Educational / Academic Development, Pedagogy or adjacent discipline. It was also possible to choose being a member of both roles (or none of them).

Depending on that role, they had to answer questions relating to their perception and awareness of the 'other' field, their professional relationship towards members of the other field, their opinion on the state of knowledge members of the other field have about the own field, their experience how much the other field's members were interested in their own field, their perceptions of distance between the two fields, and their perceptions of dynamics and trends regarding a change trend in distance between the fields. Finally, using the example of Open Educational Resources (OER) as model for a current research topic under debate, we monitored if there is a difference in opinion between the TEL and ED participants.

We only asked those participants opting for either TEL or ED (not both or other) the full set of questions. Those who had chosen 'both' or 'other', were prompted only to answer the last three questions on their perception of distance and distance trends, as well as on OER.

Overall, 28 participants replied to the survey, of which 10 participants chose the role to be TEL members, 7 chose ED and 10 chose both perspectives (1 chose 'other'). The descriptive overview with respect to the fields' distance could be drawn from the whole sample. 10 participants were female, 18 were male, and the sample was distributed over Europe with a slight over-representation of German (31%) and Swedish (10%) participants (overall 16 different countries).

Table 4. Descriptive statistics of the sample

	N	Minimum	Maximum	Mean	σ
Age	28	26	65	40.75	12.24
Distance	23	1	5	2.04	.98
Trend	25	1	3	1.8	.71
OER	28	3	5	4.69	.66

As shown in Tab. 4, the questions asked were answered on 5-point Likert scales, where the "distance between the two fields" and "trend of the distance between the fields" were coded according to 1 being no distance, 5 being a large distance, or 1 being a strongly diminishing distance and 5 being a strongly increasing distance respectively. The results show a strong agreement about the distance being low, as well as the trend of the distance being slowly diminishing. Also, regarding the value of OER (1 being the lowest and 5 the highest), a strong agreement on the positive value could be measured.

Mutual awareness and reception of the two fields showed significant differences, when splitting the sample according to the different groups into TEL and ED participants. Due to the groups being not normally distributed, we conducted a Mann-Whitney U test, with the following significant result: While the TEL ranking of awareness about ED was lower (1.9) the opposite perspective (4) was significantly higher with a Mann-Whitney U test turning out as ($U = 2.5$, $p < 0.01$).

7 Discussion

While the result of the externalized perspective of cohesion between Educational Development and TEL yield mostly descriptive results that can be interpreted as positive tendency with a bigger awareness on the ED side towards TEL, the main focus of our investigation was highlighting essential strengths of the TEL network in Europe with respect to internal cohesion between disciplines, age groups, and nationalities.

Researchers from 31 different countries participated in our TELeurope survey. All major European countries, such as Germany, UK, Spain, Italy and France, were represented. Looking at the bigger countries, especially Switzerland, Austria and the Netherlands were over-represented, while, e.g., Italy and France were under-represented.

Technology-enhanced learning has many stakeholders, such as students, lifelong learners, teachers, parents, policy-makers, decision makers in industry, and researchers [17]. The study showed, however, that there is no such thing as 'the researcher' in TEL, as disciplinary branches use different methods and practice different things. Only a detailed analysis can help to identify the needs of the different disciplinary groups.

On the other hand, the study revealed that many researchers are also engaged in policy-making. A major share of social scientists consider policy-making as part of their work in TEL, while only one third of computer scientists engage in that kind of activity. However, involvement in policy correlates with age, meaning that more experienced researchers practice those activities more often. We cannot determine from this data, if the younger computing researchers are yet to start with such activities, or if there are barriers that keep them from policy participation. As the computer scientists are strongly connected in the European network, it doesn't seem that a lack of community commitment is the reason for not doing policy-related work.

We observe that the study subject is often a significant factor in identity, as only a few researchers identify with a discipline other than the one they studied.

There are indications that some shared repertoire of practices in TEL does exist. It consists of methods such as design-based research, user-centred design, socio-technical engineering, and qualitative as well as quantitative empirical methods. Common theories include Wenger's community of practice theory, constructivist learning theory, Papert's constructionism, and Latour's actor-network theory.

Across the disciplines, an open dialogue with a broader public is present in the TEL community, especially when comparing it with the results of other open access studies. The traditional disciplinary difference that the computer sciences publish generally more, target more international, more application-oriented, and less non-professional audiences, was visible, but not as pronounced as it had been in the comparison DFG study.

For the investigated context, the TELeurope network, mutual engagement in relationships across the network was more the case for computer science researchers. However, they are not only connected to each other, but also to a lot of different clusters with persons from contrasting disciplines.

8 Conclusions

This research could indicate ways in which researchers and funders could adjust their practices to strengthen TEL as a research community, which explicitly values interdisciplinarity. For instance, highly interdisciplinary researchers (cluster D), who are already experienced with interdisciplinary research activities, can join forces and leverage already strong connections to other researchers. This is also the case for the experienced interdisciplinary developers and computer-science-based researchers (cluster F), who could further integrate non-computing researchers into their working groups. Members of these clusters could also fruitfully engage in *sharing best practices and success stories* of interdisciplinary TEL research and engage in *initiating mobility and publishing projects* to link in particularly more traditionally oriented colleagues to broaden the interdisciplinary backbone of TEL research. This is important for editors-in-chief of journals and their publishers, as well as the programme chairs of relevant conferences to react by scoping mission statements and expected topics.

One finding was that the application-oriented researchers (cluster E), who do not work in universities, are somewhat disconnected from the European TEL network. In order to increase transdisciplinary collaboration between academia and industry, funding bodies might prioritize those projects that involve partners intersecting both worlds.

It is almost certainly the case that the clusters we have defined in our study are not static, but dynamically changing. Also the TEL/ED comparison points at an increasing trend that awareness is rising especially on the Educational Development side towards TEL, but not so much from the opposite direction. Future research could measure these dynamics in greater detail, looking also at how early career researchers from education and computing grow into the community and where barriers lie. There is also a need for more research on the question of how scientists from different theoretical backgrounds view and choose their career paths, and what are reasons to leave an interdisciplinary field. Therefore motivations for interdisciplinary work and research on incentives to stay in an interdisciplinary field are very important topics for further investigation.

References

1. Del Favero, M.: Academic Disciplines. In: Guthrie, J. (ed.) Encyclopaedia of Education, 2nd edn., pp. 9–14. Macmillan, New York (2003)
2. Wild, F., Ullmann, T., Anastasiou, L., Scott, P.J., Kraker, P., Lindstadt, S., Para, G., Vandeputte, B., Duval, E., Nagel, T., Herder, E., Schweiger, S., Balacheff, N.: The STELLAR Research 2.0 Infrastructure. Deliverable D6.6, STELLAR Consortium (2012), `http://www.stellarnet.eu`
3. Strathern, M.: Interdisciplinarity: Some models from the human sciences. Interdisciplinary Science Reviews 32(2), 123–162 (2007)
4. Moran, J.: Interdisciplinarity, 2nd edn. Routledge, London (2010)

5. Voigt, C., Heinze, N., Herder, E., Cress, U.: 4th Stellar Evaluation Report – including Social Network Analysis. Deliverable D7.5, STELLAR Consortium (2011), http://www.stellarnet.eu
6. Ebner, M., Reinhardt, W.: Social networking in scientific conferences – Twitter as tool for strengthen a scientific community. In: Science 2.0 Workshop at the EC-TEL 2009 Conference, Nice, pp. 1–8 (2009)
7. Wenger, E.: Communities of Practice: Learning, Meaning, and Identity. University Press, Cambridge (1998)
8. Deutsche Forschungsgemeinschaft: Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access. Wiley-VCH, Bonn (2008)
9. Sheppard, A.G.: The sequence of factor analysis and cluster analysis: Differences in segmentation and dimensionality through the use of raw and factor scores. Tourism Analysis 1(1), 49–57 (1996)
10. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. University Press, Cambridge (1994)
11. Nowotny, H., Scott, P., Gibbons, M.: Mode 2 Revisited: The New Production of Knowledge. Minerva 41, 179–194 (2003)
12. Papert, S., Harel, I.: Situating Constructionism. In: Papert, S., Harel, I. (eds.) Constructionism. Ablex Publishing Corporation, Norwood (1991)
13. Latour, B.: Reassembling the Social: An Introduction to Actor-Network-Theory. University Press, Oxford (2005)
14. Conole, G., Scanlon, E., Mundin, P., Farrow, R.: Technology enhanced learning as a site for interdisciplinary research. TLRP TEL Programme Report, 4 (2010)
15. Kerr, A., Lorenz-Meyer, D.: Working together apart. In: Felt, U. (ed.) Knowing and Living in Academic Research, pp. 127–167. Institut of Sociology, Prague (2009)
16. Heylighen, F., Joslyn, C.: Cybernetics and second order cybernetics. Encyclopedia of Physical Science & Technology 4, 155–170 (2001)
17. Barak, N., deVries, F., Joubert, M., Specht, M., Windrum, C., Persico, D., Kraker, P.: Report of the process of establishing a Stakeholder Panel. Deliverable D5.1, STELLAR Consortium (2009), http://www.stellarnet.eu

Learning Design Studio: Educational Practice as Design Inquiry of Learning

Yishay Mor[1] and Orit Mogilevsky[2]

[1] Institute of Educational Technology, The Open University, UK
yishay.mor@open.ac.uk
[2] Technologies in Education Program, University of Haifa, Israel
oriti.mog@gmail.com

Abstract. Recently we are urged to transform education into an evidence based profession, and promote scientific standards or practice. These calls are not new – they seem to emerge every few years. We do not argue with their goal, but we contend that the suitable frame of reference is the paradigm of design science, rather than the common metaphor of medical research. This paper proposes Design Inquiry of Learning as a projection of educational design science into a professional domain, and offers the Learning Design Studio as a pedagogical manifestation of this approach. The learning design studio is a collaborative, blended, project based framework for training educators in effective and evidence-based use of educational technology. We present its theoretical underpinnings, note its fundamental principles and structures, and review three independent cases where it has been trialed. The results show that this model is effective in developing learners' theoretical knowledge as well as their practical skills, and allows them to link the two. However, it requires a considerable commitment of both learners and tutors, and may not be applicable in more casual settings.

Keywords: Learning design, teacher training, professional development, inquiry based learning, learning design studio.

1 Introduction

Recently we are urged to transform education into an evidence based profession, and promote scientific standards or practice. In fact, such calls seem to resurface every few years. The cause itself is laudable: if we wish to provide learners with effective opportunities to gain the knowledge, skills, and attitudes that they seek (or that we wish to bestow on them), and we want to do this in an efficient use of resources – we need to apply scientific rigor to our practice. The problem we see with such calls is twofold: first, they often place the onus on teachers, who are requested to adopt a more "scientific" or "research based" stance. Yet teachers typically find scientific research hard to apply in their non-research daily settings. The other, perhaps more fundamental flaw, is the implicit model of scientific knowledge and the modes of its production. While medical research is often cited as the metaphor for a desirable

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 233–245, 2013.

transformation of educational practice, we argue that the paradigm of design science is a more suitable frame of reference.

In order to instill scientific rigor as a mode of practice, we need to raise practitioners awareness to the necessity of such rigor, and equip them with the tools to support it. In other words, we need to guide practitioners in adopting and developing appropriate epistemic practices: the practices by which they establish knowledge within their domain. What are the appropriate domain-specific epistemic practices for the practical application of technology in education? Teachers operate in a complex and dynamic domain – the background knowledge and practices of their students constantly change, the technologies and resources at their disposal are perpetually evolving, and the guidance and directives they receive are frequently updated. Within this domain, they need to habitually devise new means for achieving educational goals – engendering change in their students' knowledge, behaviors, or attitudes. We posit that this is fundamentally a task of learning design, and the appropriate epistemic practice is one of design inquiry of learning. This paper presents the "Learning Design Studio", a course format aimed at enculturation educational professionals into design inquiry of learning. We note three courses and a MOOC which were based on this format, review some results, and consider their implications.

2 Background

Ben Goldacre, in a recent position paper commissioned by the UK department of education [8], called for making "teaching a truly evidence-based profession" by establishing a norm of randomized control trials. This call was answered by vocal objections in the educational research community. As Mary James notes [12], such arguments are not new. David Hargreaves [11] promoted "teaching as a research-based profession" in 1996, a position echoed by Philip Davies [8], among others. Yet, as Davies notes – the fault is not with teachers. Academic research, he argues (in agreement with Hargreaves), is often not relevant or not accessible to practitioners. Mellar, Oliver and Hadjithoma-Garstka [18] find that research is perceived by practitioners as providing too much detail, or conflicting evidence, does not address their immediate concerns or does not acknowledge the reality of their experiences. Ironically, they conclude, "the same characteristics that make it hard to draw general principles from the work can also make it credible to practitioners". Korthagen et al [15] show that teacher training which focuses on educational theory fails consistently. Not only do teachers find themselves ill-equipped to translate the theoretical abstractions to the concrete context in which they work, their negative experience in attempting to do so results in theory aversion: teachers feel threatened by educational theory and see teacher education as detached and useless.

This sense of dissonance between educational research and practice often leads practitioners and policy makers to disengage with research, seeing teaching as "a craft and it is best learnt as an apprentice observing a master craftsman or -woman." [9, in 3]. Yet, as [23] show, the most successful educational systems are those maintain multi-directional links between research, practice and teacher training.

We hold that the weakness in the position of Goldacre and others before him is not in the aspiration for scientific rigor, but in the scientific frame of reference. While both Goldacre and Hargreaves draw their examples from positivist research in medicine, Diana Laurillard [16] argues that teaching should be repositioned as a design science. Laurillard adopts Simon's [26] paradigmatic distinction between natural science which describes how the world is, and design science which is concerned with how it should be. Design science is value laden, functionally oriented and representation sensitive [22]. Design, in this context, is the informed creative practice of devising "*courses of action aimed at changing existing situations into desired ones*" [26, p 129]. In identifying certain situations as more desirable than others, design science starts off from a value proposition. It seeks to promote change, rather than describe the current state of affairs, and is thus focused on function rather than structure. Finally, since it deals with human action and perception – the manner in which ideas are (re)presented is important, not just their validity.

One potent emerging approach to linking research and practice is *Teacher Inquiry into Student Learning (TISL)* [5]. Following an action research tradition, the TISL approach addresses the professional development of teachers by investigating student learning through action-oriented, evidence-based inquiry.

Another emerging approach addresses teachers' professional development by positioning them as learning designers and focusing on the practical process of devising effective learning experiences [13]. Recent studies demonstrate how training teachers as learning designers enhances not only their practical skills but also their theoretical understanding [4], [16], [25], [26].

We propose a synergy of these approaches: projecting the idea of teacher inquiry onto the paradigm of design science leads us to a model of teaching a *Design Inquiry of Learning (DIL)*. Inquiry-based learning attempts to shape educational experiences in the model of scientific investigation. Similarly, an inquiry approach to the training of educational practitioners should mimic the form of design research in education.

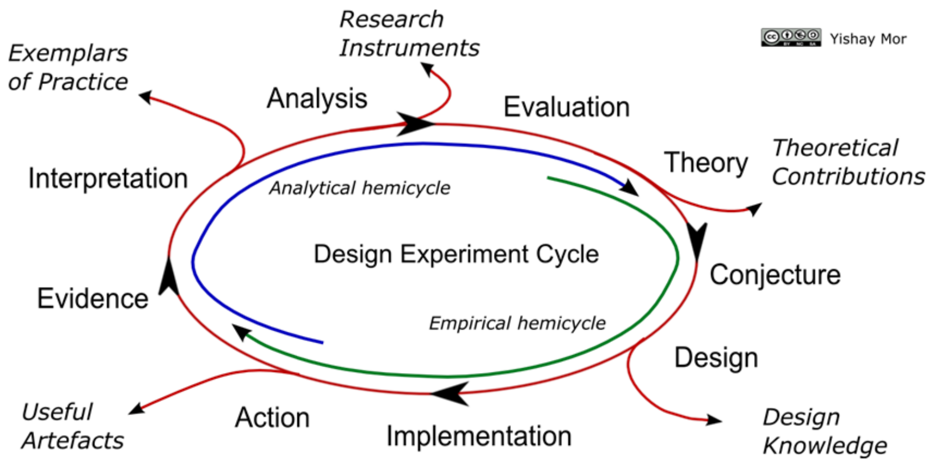


Fig. 1. The Design Research Cycle

Thus, the DIL approach mimics the structure of an educational design study [22], with the exception that students do not have the resources or the time to conduct several iterations, scaling up from a conceptual prototype to an extensive deployment.

Design based research progresses through cycles of theoretical analysis, conjectures, design, implementation, analysis and evaluation – which feeds into adjusting the theory and deriving practical artefacts (Fig. 1) [18]. Anastopoulou et al [1] describe personal inquiry learning as a cycle of questioning, investigation, evidence collection, analysis, sharing, and reflection. Combining these two yields the cycle of design inquiry of learning: imagining a desired change, investigating the current situation, drawing inspiration from theoretical frameworks and exemplars of practice, ideating and designing an innovation, prototyping it, evaluating its effects and reflecting on the process.

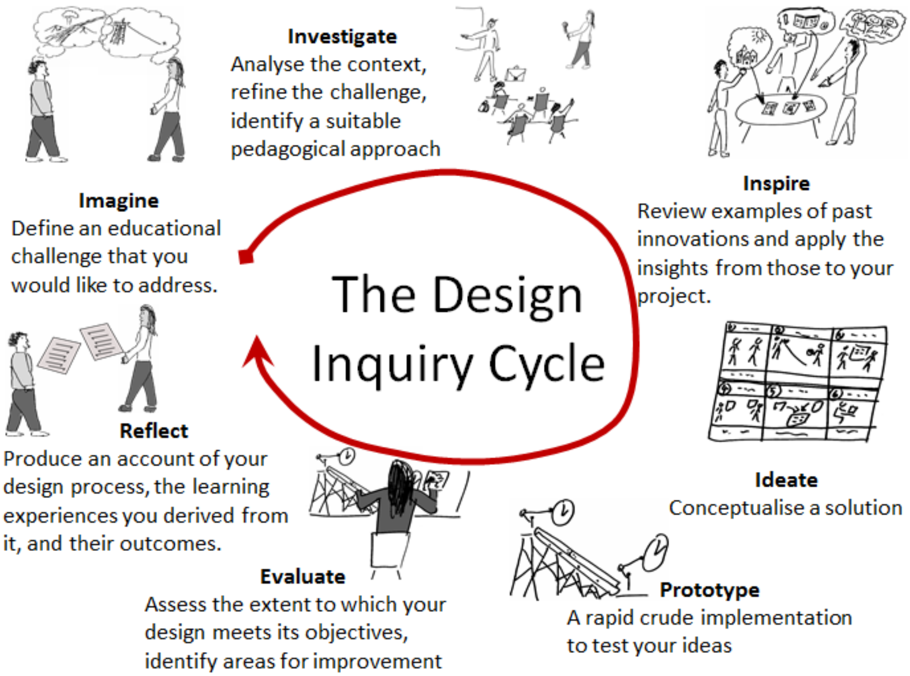


Fig. 2. The Design Inquiry of Learning Cycle

3 The Learning Design Studio

One approach which appears to hold significant promise in training learning designers is the learning design studio (LDS) [13], [4]. This approach is modelled after the tradition of studio-instruction in arts and design disciplines (such as architecture). In this model, the main activity of a course is the students' continued work on design challenges in a defined domain of practice. Students typically work in groups. They identify an educational challenge, research it, and devise innovative means of addressing it. The course instructor guides the students through the process, and classroom sessions are mostly dedicated to group work and public review of design artefacts.

The model of learning design studio presented in this paper manifests the DIL approach presented above.

In a LDS, students work in groups on projects of their own choice. Each group identifies a concrete educational context and a specific educational challenge within this context, locates and reviews relevant literature, devise a techno-pedagogical innovation to address the chosen challenge in its context, and evaluate their innovation – if possible, by observing its implementation in the real-world context.

The first phase of an LDS course focuses on defining the context in which projects will be situated and the pedagogical challenge they attempt to address within this context. Students are asked to propose an idea for a project they would like to develop. They form groups based on common interests, and spend the majority of the course time working on their joint project. Students document and described the material, social and intentional factors which define the environment in which they will work. Reflecting on the tensions identified in the analysis of the context, students are asked to specify well-defined and measurable educational objectives. Next, they conduct preliminary research, reviewing appropriate learning theories and relevant case studies, and choosing the theories which they identify with and the cases which inspire them, as a basis for their design work.

Based on their articulation of the context and challenge, and the outcomes of their preliminary research, students develop an initial scenario, which included an outline of the proposed solution, and a storyboard depicting the learner's envisioned activities and expected learning trajectory. Students consult existing repositories of design knowledge, such as the design principles database [14] or appropriate collections of design patterns [18], [22]. They articulate the knowledge they gathered in the form of a prototype of their solution. This prototype is evaluated, if possible – through a pilot study in the actual project settings, and if not – via a heuristic evaluation by peers or experts.

Despite the seemingly pragmatic, action-oriented nature of the LDS, individual and group reflection are central to the learning experience. Students are instructed to maintain a learning journal, provide peer feedback within and between project groups, and conclude the process by writing a design narrative, recounting their journey.

Students use a website as a collaborative workspace which scaffolds them through the LDS process. When they complete their project, they edit their website to present their work - the design process, its outputs, and their reflections.

4 Implementation

The learning design studio format was trialled in two courses at the Technologies in Education postgraduate programme at the University of Haifa during the academic year 2010-2011, in the Open Learning Design Massive Open Online Course (OLDS

MOOC) and in one course in The Open University's Master in Open and Distance education programme.

The University of Haifa cases were a course on "games and learning" (http://courses.edtech.haifa.ac.il/games) and a course on "mobile learning" (http://courses.edtech.haifa.ac.il/mlearning) [21]. Both ran for 13 weeks in a blended format (2 hours face time, 4 hours independent study). The first included 22 students, who split into 9 project groups. The second included 17 students in 6 project groups. The courses used the institutional google apps suite as a platform. Students created a project site from a template which was provided, and used it throughout the course. The website template contained sections corresponding to the phases of a single iteration of a design experiment. Students replaced the instructions in the template with the content and artefacts they generated in the course of their work, so that when they completed the project, the website presented both its products and the process by which they were created. All students passed, and all projects were completed.

The OLDS MOOC (http://olds.ac.uk) ran for nine weeks, from Jan. to March 2013. The course was designed for 3-10 hours a week, but many participants could not commit to this timeframe. Out of the 2420 who registered their interest in the course, 200-300 where still following the course in week 8, but only 23 where actively contributing to the course space. The course used a google site and two google groups as its focal space, with "clouds" and "cloudscapes" in cloudworks (http://www.cloudworks.ac.uk) to support specific activities. Some participants used their personal blog as their medium for participation.

The OU case was a 7 week block out of the 30 week MA course "openness and innovation in elearning" (http://www3.open.ac.uk/study/postgraduate/course/h817.htm). This course has 70 students registered. It is taught fully online and students are expected to commit 14 hours a week. The students were assigned to 11 project groups

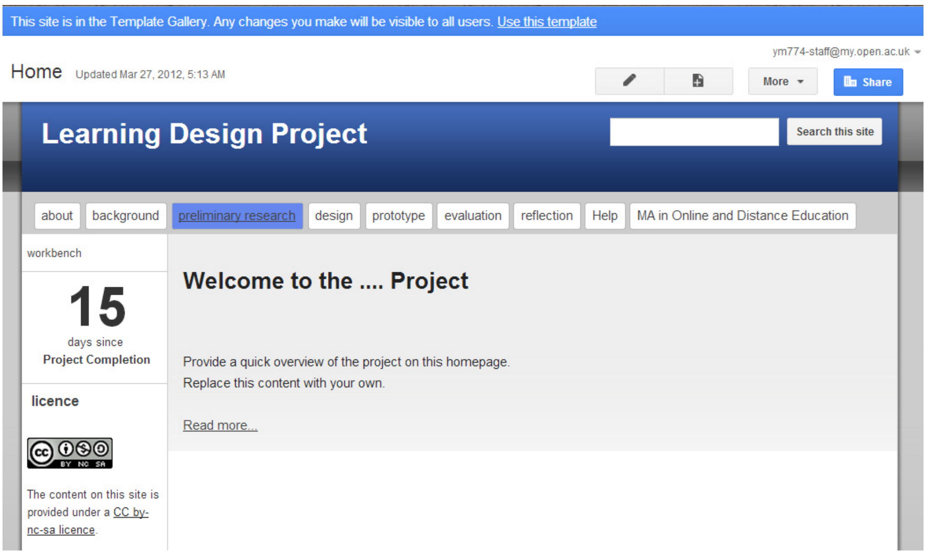


Fig. 3. OU project website template

based on their choice of project subject. The course used the OU version of Moodle as a VLE, and students also used the institutional suite of google apps as their collaborative workspace. Similar to the University of Haifa courses, students were provided with a website template (Fig. 3) which scaffolded their work. This template offered website sections which corresponded to the various phases in the project lifecycle. Students edited and populated this site as they progressed.

In addition, the course used a bespoke system called OpenDesignStudio for sharing and discussing their work between project groups. 10 out of the 11 projects were completed successfully, most with impressive outputs. All projects received a large number of comments from peers in other project groups. This block has ended very recently, and is currently being evaluated. As in the University of Haifa courses, each group maintained a website for their project, instantiated from a template designed to scaffold their design process.

4.1 Variations

The characteristics of the three domains of implementation were radically different, and consequently the design of the LDS, and the supporting technology, had to be adjusted significantly to suit the different situations.

The University of Haifa courses where conducted in a blended environment: weekly classroom sessions with continuous interaction between students and tutor via the course forum, email, and skype. The classroom sessions were dedicated mainly to summarizing the online discussions and reviewing students' progress on their projects. Students often met out of classroom to work on their projects together – either physically or virtually (e.g. by skype). The relatively longer time span of the courses (13 weeks) meant that most students had the opportunity to test their design in real-world conditions, with learners representing the target audience, adding a great deal of depth to their learning experience. The project sites were shared in a "walled garden" during the course – visible only to the course community. After completion, students had the option of making their projects public, and most chose to do so.

The OLDS MOOC assumed a radically open stance: no registration was required, all resources were, and still are, freely and openly available, and facilitators repeatedly noted that peripheral, casual and incidental participation were legitimate. Consequently, participants found it hard to form project teams. The open nature of the course also implied that we did not have the option of setting up a project site template for participants to use. The majority of participants found it hard to form project teams, and did not systematically work through the proposed phases of a learning design project. These limitations, to an extent, were predicted by the course team. In response, the tasks were designed as autonomous building blocks, which can be experienced individually but can also stack up to more complex and rich learning structures. The open nature of the course also meant that participants work was typically public throughout the course and beyond, although it also meant that they had an option of keeping their work completely private. Another significant difference was that the MOOC was facilitated completely online, and the ratio of facilitators to participants was two scales of magnitude higher. This made the classical studio format irrelevant. To provide an alternative, the course team held weekly "fishbowl" sessions, where a few facilitators and a few participants conducted an hour long video conference, discussing the weeks' tasks and the participants' projects. Other participants could watch the conference live and interact via twitter or view a recording later.

The context of the OU course was, in many ways, half-way between the small-scale blended format of the University of Haifa and the large-scale on-line format of the MOOC. The number of student was 3-4 times larger than the Haifa course, but the ratio of students to facilitator was better. The time span was much shorter, but students were expected to dedicate more hours a week. The course was conducted completely online, but with a rich set of supporting technologies. Consequently, the studio format was maintained, albeit mediated by forums and the OpenDesignStudio rather than face to face. The shorter time span and the geographical distribution of students made real-world user testing unfeasible. Instead, students were guided in conducting a heuristic evaluation of peer projects.

5 Results

5.1 University of Haifa

The University of Haifa courses have been evaluated in detail in [21]. We recount the main findings here.

All 39 students completed the courses successfully, and responses to the end of course survey suggest they valued its contribution to their understanding of the core issues presented, as well as the pragmatic considerations of implementing these ideas in realistic educational contexts. Students expressed notable criticism about the courses' administrative aspects, as well as the workload which exceeded their expectations. Despite these shortcomings, students all acknowledged the effectiveness of the design studio approach, some noting that it has changed their attitude to the course subjects, and to technology enhanced education in general. Content analysis of the mobile learning course students' design narratives and learning journals uncovers several themes:

- Students initially found the design-inquiry approach confusing, and engaged with it at a superficial level. However, in retrospect students acknowledged the advantages of the design-inquiry approach. To an extent, the initial confusion was alleviated by the iterative dynamics of the design-inquiry process.
- Students reported on their difficulty in concretisation of theories and abstract ideas. The fact that students reflect on this issue indicates that they are aware of it, and indeed – some of their comments suggest the process helped them take steps to address it.
- The design inquiry process at the centre of this course was supported by a variety of tools, methods and representations: a project site template, a design scenario template, force maps, design principles, storyboards, etc. Indeed, students acknowledged the value of these tools.

- Classroom sessions focused on guided group work, and groups' presentation of their progress to the course assembly. Students commented on the contribution of these interactions to their learning.

The ultimate test of a learning intervention is in its sustained effects on learners – their knowledge, attitudes and practices. Paradoxically, university courses are typically evaluated shortly after their completion, thus measuring only their short-term effects. In an attempt to counter this observation, we surveyed the students at the University of Haifa courses two years after the course completion. Using the email addresses students had used at the time of the course, we circulated a short web-based survey. This survey consisted of 9 likert-type questions, and one open comment. The questions aimed to assess students' perception of the course's contribution to their theoretical knowledge and their professional practices. 16 out of the 38 eligible students (excluding the author of this paper) responded to the survey. Fig. 4 shows a summary of the responses. These suggest that even with the advantage of hindsight, students acknowledge both the theoretical and the practical contribution of the course.

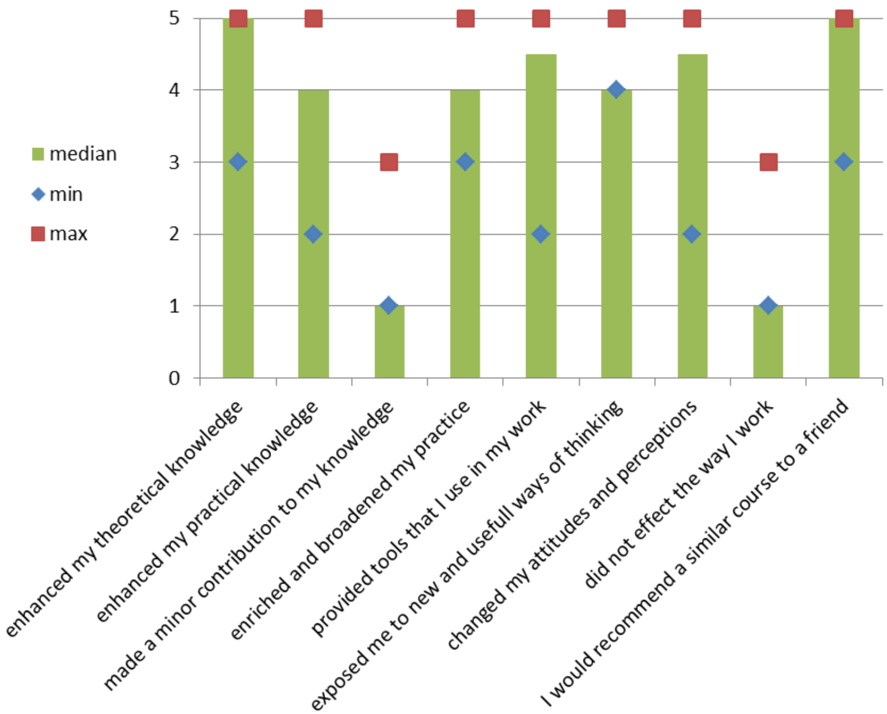


Fig. 4. Results University of Haifa long-term effects survey

5.2 OLDS MOOC

Overall, the OLDS MOOC was fairly successful in achieving its aims. It raised awareness to the field of learning design, and exposed a broad audience to a variety of tools, techniques, and methodologies. The MOOC resources are open and freely available, and are frequently used: the last week saw approximately 700 visitors on the MOOC site. We have anecdotal evidence from participants testifying that the MOOC has changed their professional perceptions and practices.

Nevertheless, as [6] notes: “Collaborative group working as implemented in this course design did not achieve the desired result. Consequently, whilst a good number of participants attempted to, or indeed, formed, nascent course development or study support groups these in general did not last more than a few weeks. Participants did however value the sharing, commenting and feedback that too place in broader course space and the sense of community this engendered”

Likewise, we are aware of only a very small number of projects (group or individual) which succeeded in following the full design inquiry of learning cycle. However, the evidence suggests that many participants did appreciate the DIL approach, and adopted elements of it, such as investigating the context of design, reviewing examples of past innovation and projecting the lessons from these to their project, ideation, prototyping and evaluation.

Taken together, the evidence suggests that the individual activities were successful, and learners found the community experience rewarding, but the Learning Design Studio approach failed. This can, to an extent, be explained by shortcomings of the supporting technology. However, even if the technical issues would have been resolved, it is questionable how many participants would have chosen to, or would have the capacity to, dedicate the sustained effort required to complete a project. It appears that the LDS format relies on strong group cohesion and intensive tutor support – both of which are hard to achieve in a MOOC setting.

5.3 OU

As noted above, the OU course LDS has only recently completed its first presentation, and it is currently being evaluated. Nevertheless, initial analysis indicates that by and large this was a success. Ten projects have completed a full cycle of investigating the project context, reviewing past cases and relevant theories, applying the lessons from these to the design and development of a prototype, and conducting a structured evaluation of the prototype.

Student reflections suggest that most of them found the LDS block challenging but rewarding, and believe they will make use of the knowledge they acquired.

However, students reported numerous technical difficulties, and noted that the extensive workload prevented them from reflecting on the process. In order to support the students in a distance learning environment, we attempted to provide very detailed instructions and scaffold activities by templates designed for the various tasks. Many students found these too prescriptive and felt that at times the tasks were reduced to form filling exercises.

Most of these issues would appear to be fairly superficial, and can be resolved by refining the guidance materials and supporting technologies. Overall, this experiment seems to suggest that the LDS format is not limited to traditional, small class, face to face scenarios – but can also be applied effectively in a large scale distance education setting.

6 Discussion

The three cases presented above suggest that the DIL approach offers significant value for educational practice, and a robust framework for training and development of educational professionals. Specifically, the manifestation of this approach in the LDS emerges as an effective format for courses in TEL. However, this format has its limitations, and may not be suitable for certain circumstances.

DIL combines an inquiry-based pedagogy with a design-based epistemology. The inquiry learning model posits that learning is more effective when grounded in active exploration of questions which are meaningful to the learners. In the cases presented here, the learners are educational professionals wishing to make effective use of technology in their practice. Thus, their inquiry explored questions pertaining to this domain. Design, in these cases, was adopted both as a mode of learning and as a mode of action: students learned *by* design, and they learned *to* design. The evaluation of the projects in the University of Haifa courses [14] showed that students had assimilated both relevant case studies and techno-pedagogical theories, by binding them to a personally meaningful context of action.

The LDS model adds the constuctionist, project-based, collaborative pedagogical features of the studio-based educational tradition of design practices. The initial feedback from the OU course suggests that students experienced a sense of satisfaction in completing a meaningful project – such a sense is arguable significant in terms of the affective dimension of learning. Students in all three cases reported that they valued the social interaction and in the two university course cases – the positive impact of the project group on their learning experience. This is particularly notable in the case of distance learning, as the OU internal surveys persistently show students aversion to collaborative learning. In both the University of Haifa and the OU cases, students were highly active in social learning interactions within and between groups. The feedbacks between groups and sharing the opinions in class and online discussions helped the students in constructing new meanings and juxtaposition different points of view.

Educational practice, and consequently – educational technology – is always context-dependent. The LDS approach places an emphasis on documenting and articulating the context in which the pedagogical challenge is situated, and carefully referring to that context in the design of the solution. As discussed in [20] – this challenge of getting students to analyse context is far from resolved, but the courses showed promising signs in this respect. Several participant comments from both the MOOC and the OU course stress this particular aspect.

As our experience suggests, and the student feedback confirms, the LDS model is a demanding one, both for students and for tutors. Its success is likely to be limited without serious commitment of both. For this reason, it is probably less suitable for more "casual" learning situations, such as MOOCs. However, where the commitment is present – it is highly rewarded.

7 Conclusions

We opened this paper, perhaps somewhat provocatively, with Ben Goldacre's call for turning education into an "evidence based profession". We argued that this is a laudable cause, but its fault lies in its implicit model of evidence, or scientific process. We presented design science as an alternative paradigm, which we claim is more suitable for educational research, and offered Design Inquiry of Learning (DIL) as a projection of that paradigm into the domain professional practice.

The main bulk of this paper reviewed the Learning Design Studio (LDS) format, as a pedagogical structure for training educational professionals in DIL. We presented its theoretical underpinnings, its fundamental principles, and a proposed sequence of learning activities. We then reviewed three recent implementations of this model.

The preliminary evidence from these three cases seems to indicate that the model has merit, and appears to deliver on its promises. Students have achieved impressive results in the course of their project work. They report that the courses have enriched them both theoretically and in terms of practical skills. Perhaps even more important, the courses manage to link the two together, and promote a critical, informed, systematic and context-sensitive mode of practice.

Technology Enhanced Learning has established itself as a significant field of scientific inquiry. Yet this field often finds it challenging to disseminate its knowledge beyond the academic circle. A DIL approach and the LDS model in particular, may offer a possible way of breaching the divide between research and practice in education.

As the OLDS MOOC example illustrates, while the DIL approach has a wide remit – the LDS model imposes considerable requirements on the learning situation, and thus may not be suitable in certain conditions.

Finally a caveat is due. The cases reported here are still but a small sample, and their accomplishments and shortcomings may be partially attributed to personal style, extraneous factors, or a simple Hawthorne effect. Nevertheless, the evidence is strong enough to warrant further independent trials of the models presented here and their implementation modes.

References

1. Anastopoulou, S., Sharples, M., Ainsworth, S., Crook, C., O'Malley, C., Wright, M.: Creating Personal Meaning through Technology-Supported Science Inquiry Learning across Formal and Informal Settings. International Journal of Science Education 34(2), 251–273 (2012)
2. Beetham, H., Sharpe, R.: Rethinking Pedagogy for a Digital Age: Designing for 21st Century Learning. Routledge (2013)
3. BERA: Research and Teacher Education: The BERA Inquiry (2013)
4. Cox, C., Harrison, S., Hoadley, C.: Applying the studio model" to learning technology design. Educating Learning Technology Designers: Guiding and Inspiring Creators of Innovative Educational Tools 145 (2008)
5. Clark, W., Luckin, R., Jewitt, C.: Deliverable D5.1 Methods and Specifications for TISL Components V1. NEXT-TELL Consortium, EU (2011)
6. Cross, S.: Evaluation of the OLDS MOOC curriculum design course: Participant perspectives, expectations and experiences, Technical report, the Open University (2013)

7. Cross, S., Conole, G., Clark, P., Brasher, A., Weller, M.: Mapping a landscape of learning design: Identifying key trends in current practice at the Open University. In: 2008 European LAMS Conference (2008)
8. Davies, P.: What is evidence-based education? British Journal of Educational Studies 47(2), 108–121 (1999)
9. DfE: The Importance of Teaching (White Paper). DfE, London (2010)
10. Goldacre, B.: Building Evidence into Education. DfE (2013), https://www.gov.uk/government/news/building-evidence-into-education
11. Hargreaves, D.: Teaching as a research-based profession: Possibilities and prospects (The teacher training agency lecture 1996). Educational Research and Evidence-based Practice, pp. 3–17 (2007)
12. James, M.: New (or not new) directions in evidence-based practice in education (2013)
13. Kali, Y., Ronen-Fuhrmann, T.: Teaching to design educational technologies. International Journal of Learning Technology 6(1), 4–23 (2011)
14. Kali, Y.: Collaborative knowledge building using a design principles database. IJCSCL 1, 187–201 (2006)
15. Korthagen, F.A.J., Kessels, J., Koster, B., Lagerwerf, B., Wubbels, T.: Linking practice and theory: The pedagogy of realistic teacher education. Lawrence Erlbaum (2001)
16. Laurillard, D.: The teacher as action researcher: Using technology to capture pedagogic form. Studies in Higher Education 33, 139–154 (2008)
17. Laurillard, D.: Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology (Paperback) - Routledge. Routledge (2012)
18. Mellar, H., Oliver, M., Hadjithoma-Garstka, C.: The role of research in institutional transformation, . In: Mayes, T., Bullen, P., Mellar, H., Oliver, M. (eds.) Transforming Higher Education through Technology-Enhanced Learning. Higher Education Academy, York (2009)
19. Mor, Y.: SNaP! Re-using, sharing and communicating designs and design knowledge using Scenarios, Narratives and Patterns. In: Luckin, R., Goodyear, P., Grabowski, B., Puntambekar, S., Winters, N., Underwood, J. (eds.) Handbook of Design in Educational Technology. Routledge (2013)
20. Mor, Y.: Context is what we take for granted: addressing context in design-centric teacher training (2011)
21. Mor, Y., Mogilevsky, O.: The Learning Design Studio: Collaborative Design Inquiry as Teachers' Professional Development, Research in Learning Technology (forthcoming)
22. Mor, Y., Winters, N.: Design approaches in technology enhanced learning. Interactive Learning Environments 15, 61–75 (2007)
23. OECD: Lessons from PISA for the United States: strong performers and successful reformers in education (2011), http://www.oecd.org/dataoecd/32/50/46623978.pdf
24. Ronen-Fuhrmann, T., Kali, Y.: The role of concretization in acquiring design knowledge. In: Proceedings of the 9th International Conference of the Learning Science, vol. 1, pp. 468–475 (2010)
25. Ronen-Fuhrmann, T., Kali, Y., Hoadley, C.: Helping Education Students Understand Learning Through Designing. Educational Technology 48, 26–33 (2008)
26. Simon, H.A.: The Sciences of the Artificial, 3rd edn. The MIT Press, Cambridge (1996)
27. Voogt, J., Westbroek, H., Handelzalts, A., Walraven, A., McKenney, S., Pieters, J., De Vries, B.: Teacher learning in collaborative curriculum design. Teaching and Teacher Education 27(8), 1235–1244 (2011)

Usage Context-Boosted Filtering for Recommender Systems in TEL

Katja Niemann and Martin Wolpers

Fraunhofer Institute for Applied Information Technology FIT
Schloss Birlinghoven, 53754 Sankt Augustin, Germany
{katja.niemann,martin.wolpers}@fit.fraunhofer.de

Abstract. In this paper, we introduce a new way of detecting semantic similarities between learning objects by analysing their usage in web portals. Our approach does not rely on the content of the learning objects or on the relations between users and learning objects but on the usage-based relations between the objects themselves. We take this new semantic similarity measure to enhance existing recommendation approaches for use in technology enhanced learning.

Keywords: item similarity, recommender systems, usage context.

1 Introduction

In recent years, recommender systems have become popular in the Technology Enhanced Learning (TEL) domain to identify suitable learning resources for users e.g. offered by learning portals. Such recommender systems often require extensive additional information about the learning objects, e.g. the competencies / knowledge they impart and information about the users, e.g. existing knowledge and learning goals [1]. Such additional information is often not available and expensive to create, thus, we need new possibilities to gather informations about users and items that are implicitly given in their behaviour and usage, respectively.

The special interest group dataTEL[1] (Data-driven Research and Learning Analytics) was created to increase research on educational datasets and make educational systems more transparent and predictable. In a first study of the datasets submitted to dataTEL, Verbert et al. [2] found that challenges to be tackled include sparsity of data and that further research on implicit relevance indicators and similarity measures is required to find relevant items and/or users.

In this paper, we present a new way of detecting similarities between learning objects by considering their usage contexts, i.e. the learning objects they most significantly often co-occur with in the same sessions and claim that usage context-based similarity gives rise to content similarity and can thus be used for recommendations. This way, we are able to compensate a lack of semantic information for the learning objects as well as sparsity of rating data. We evaluate

[1] http://ea-tel.eu/sig-datatel/

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 246–259, 2013.

our approach using the usage and semantic metadata collected in the MACE[2] and in the TravelWell[3] portal.

The paper is structured as follows: In chapter 2, we give an overview of existing recommendation strategies and in chapter 3, we describe the principles of the usage context-based approach we propose. In chapter 4, we present the datasets MACE and TravelWell that are used in the experimental evaluations presented in chapter 5 (usage context-based similarity vs. content-based similarity) and chapter 6 (usage context-based similarity for enhancing existing recommender systems). In chapter 7 we give a conclusion and present ideas for further work.

2 Related Work: Recommendation Strategies

2.1 Content-Based Filtering

Content-based systems use the items' attributes and the users' preferences for recommendations. Item profiles can be created automatically, e.g. through keyword extraction for text documents, or manually, e.g. for restaurants holding attributes like style and location. User profiles can be built explicitly by asking the users about their interests or implicitly by the users' given ratings. For recommending, the user profiles are matched against the item profiles and the most suitable unknown items are recommended. Several systems are developed that use content-based filtering to help users find information, e.g. PRES (Personalized Recommender System) [3] and Syskill & Webert [4]. Recommender systems relying on content-based filtering suffer from problems like the new user problem and overspecialisation. Additionally, it can be time-consuming and expensive to maintain the item profiles.

2.2 Collaborative Filtering

Systems based on collaborative filtering do not consider the item's attributes but make use of user ratings on items that can be explicit (e.g. rate a book with 3 stars) or implicit (e.g. visit a site, listen to a song). In user-based collaborative filtering an item is suggested to a user based on ratings of that item by users most similar to her. A prominent example for such a system is MovieLens[4]. Item-based collaborative filtering approaches do not compare users but calculate the similarity of items by comparing their users' implicit and explicit ratings [5]. A system relying on this approach is Amazon.com. where products often bought by the same users get a higher similarity value than products that do not share so many users [6]. Advantages of collaborative filtering are that no semantic information is needed to create item profiles and cross-genre niches can be identified. However, user and item profiles first need to evolve before sufficient recommendations can be produced.

[2] http://mace-project.eu/

[3] http://lreforschools.eun.org/web/guest/travelwell-all

[4] http://movielens.umn.edu

2.3 Matrix Factorization

Matrix factorization models map both users and items to a joint latent factor space, i.e. each item is associated with a vector q that measures the extent to which the item bears the factors of the space and each user is associated with a vector p that measures the extent of interest the user has in items that are high on the corresponding factors. The resulting dot product of p and q represents the users overall interest in the items characteristics [7]. Matrix factorization methods have become popular since they combine a high predictive accuracy with good scalability. In recent years, several approaches have been created to deal with the major challenge of matrix factorization, which is computing the mapping of each item and user to factor vectors, see [8] and [9]. Similarly to collaborative filtering, the more ratings are given for users and items, the more accurate are the predicted ratings.

2.4 Hybrid Systems

Hybrid systems are implemented to exert the advantages from more than one technique while the drawbacks of single techniques can be compensated. Burke [10] describes several approaches to combine recommenders, e.g. Ranked Hybrid (combining the ratings from different recommenders using a weighting scheme, see [11]), Switching Hybrid (selecting the rating prediction from the recommender with the highest confidence value, see [12]), and Feature Augmentation Hybrid (each contributing recommender adds features to the items' descriptions, so that the actual recommender has a better base, see [13]).

3 Usage Context-Based Similarity

3.1 Background of the Notion Usage Context

In the following, we will focus on the notion of usage context that was inspired by the concept of word contexts successfully used in linguistics. Words stand in linear orders, e.g. in speech or in written texts. The context of a word can thus easily be defined by the words that occur before and after it. If two words have very similar contexts (e.g. words they often co-occur with in sentences), they are said to be paradigmatically related [14]. For example: In many contexts, the word "car" can be replaced by the word "vehicle", i.e. they share a similar context containing e.g. the words "driver" and "highway". Thus, paradigmatic relations lead us to semantic relations or, as [15] states, context similarity correlates with content relatedness. We take up this insight and form our hypothesis that it holds true not only for words but also for other entities that usage context similarity provides an indication for content relatedness.

Similar to words used in sentences, data objects are used in sessions. The definition of the term "session" depends on the conditions. For example, a session can comprise all events conducted by a user between logging in and out of a portal. Thus, the usage context of an object is the result of the objects accessed before or after it in the same sessions, i.e. by its (significant) co-occurrences. When comparing two data objects, they are assumed to be semantically related, if their usage contexts' significantly overlap. Please note that two objects can have highly similar usage context, even if they were never used together.

3.2 Defining Significant Co-occurrences

We define two objects to be co-occurrences if they co-occurr in at least one session. However, not every co-occurrence is significant, rather most co-occurrences are coincidental. Basic association measures calculate a significance score by comparing the observed frequency O of a co-occurrence with its expected frequency E, examples are MI (mutual information, formula 1) and z-scores, see [16]. These simple association measures often give close approximation to the more sophisticated association measures (as described below) and are therefore sufficient for many applications. They also have some limitations as they, for instance, tend to fail when calculating the significance value for a very frequent and an infrequent object [17].

$$\text{MI} = log_2 \frac{O}{E} \tag{1}$$

In statistical theory, association measures and independence tests are always based on a cross-classification of a set of objects, e.g. using contingency tables. These measures compare the expected and the observed frequencies as well. In contrast to the more simple approaches, they do not only consider the expected co-occurrence frequency of the two objects but compute the expected frequencies for all cells in the contingency table [17]. Table 1 shows the contingency table for the objects i and j which co-occured O_{11} times. Additionally, i was accessed in O_{12} sessions in which j was not accessed, j was accessed in O_{21} sessions in which i was not accessed and O_{22} session use any of these s at all. The expected values for these observed values are E_{11}, E_{12}, E_{21}, and E_{22}, respectively.

Table 1. Contingency table

	j	$\neg j$	
i	O_{11}	O_{12}	$=R_1$
$\neg i$	O_{21}	O_{22}	$=R_2$
	$=C_1$	$=C_2$	$=N$

Commonly used association measures that are based on contingency tables are the chi-squared test and log-likelihood, see [16]. The χ^2 test adds up the squared z-scores for each cell in the contingency table and puts them in relation to the expected frequencies. Since the normal approximation implicit in the z-scores becomes inaccurate if any of the expected frequencies is small [17], the Yates' continuity correction [18] shown in equation 2 offers a better approximation (corrected χ^2-test). Equation 3 shows the log-likelihood measure [19].

$$\chi^2\text{-corr} = \frac{(|O_{11}O_{22} - O_{12}O_{21}| - \frac{N}{2})^2}{R_1R_2C_1C_2} \quad (2)$$

$$\text{log-likelihood} = 2\sum_{ij} O_{ij} ln \frac{O_{ij}}{E_{ij}} \quad (3)$$

After the calculation of the co-occurrences' significance values, the most significant ones must be selected for each object. There are two ways to do so, i.e. by ranking or by using a threshold. Ranking means that the co-occurrences are sorted by their significance values and only the n most significant co-occurrences are selected. When using a threshold, only co-occurrences with a significance value higher than the threshold are selected. However, there is no standard scale of measurement to draw a clear distinction between significant and non-significant co-occurrences [20]. Therefore the calculation of a suitable n or a suitable threshold (depending on the approach) is an exploratory investigation.

3.3 Object Similarity Calculation

We calculate the similarity for each object pair using the cosine similarity where each object i is described by a vector which holds the most significant co-occurrences of object i. The cosine similarity measures the angle between two vectors, thus, the significance values of the co-occurrences are considered.

3.4 Expected Rating Calculation

We compute the expected rating $p(u, i)$ on an object i for a user u by averaging the ratings $r(u, i)$ given by the user to the other objects in her profile $P(u)$ while each rating is weighted by the corresponding similarity $sim(i, j)$, see equation 4.

$$p(u, i) = \frac{\sum_{j \in P(u), i \neq j} (sim(i, j) * r(u, j))}{\sum_{j \in P(u), i \neq j} |sim(i, j)|} \quad (4)$$

4 Datasets

4.1 MACE

The MACE (Metadata for Architectural Contents in Europe) project relates digital learning resources about architecture, stored in various repositories, with each other across repository boundaries to enable new ways of finding relevant information [21]. While interacting with the MACE portal, users are monitored and their activities are recorded as CAM (Contextualized Attention Metadata, [22]) instances. The event types considered for creating the sessions are accessing the metadata of a learning object in the MACE portal, e.g. its ratings or user tags and accessing the learning object in its origin repository. All other events that involve learning objects, e.g. tagging and rating, require the access of the learning object's metadata, thus, the learning object is already part of the session without considering these events. Each CAM instance comprises at least the event type, the identifier of the user who conducted the event, a timestamp, and the identifier of the involved object. The CAM instances used for the evaluation were collected in a period of three years from September 2009 to October 2012. Overall, we considered CAM events for 12,176 learning objects conducted by 620 users in 4,291 Sessions. A session comprises on average 6.28 distinct learning objects and each learning object is used in on average 2.18 sessions.

MACE offers users and domain experts the possibility of editing parts of the metadata, e.g. tags and classifications. We use the tags and classifications to create semantic similarities between learning objects as baselines to compare our results to. The tags are free text and can be assigned to learning objects by logged in users. The classifications are defined in a controlled vocabulary consisting of 2884 terms and can only be set by domain experts. 78.69% of the used learning objects hold such additional semantic metadata, 70.8% hold tags, 14.83% hold classifications, and 8.82% hold both. Each tagged learning object holds on average 6.59 tags and each classified learning object holds on average 2.27 classifications. Additionally, logged in users are able to rate learning objects. We use the provided ratings to test and evaluate our recommender systems. In total, 230 learning objects were rated by 73 users, each of these learning objects was rated at least once and at maximum 4 times (on average 1.2 times), and each of the 73 users rated 1-19 learning objects (on average 3.79). This results in a user-item rating matrix with a sparsity of 98.35%.

4.2 TravelWell

The TravelWell dataset [23] was collected on the Learning Resource Exchange (LRE) portal that makes open educational resources available from more than 20 content providers in Europe and elsewhere. The dataset contains information about the rating and tagging behaviour of 98 registered users over a period of six months (August 2008 - February 2009). For each user activity, the date, user identifier, identifier and the tag, respectively the rating is stored. As there is no timestamp but only the date, a session comprises all activities conducted by a user in one day.

Overall, 14,248 events took place in 255 sessions where each session comprises 55 distinct learning objects on average. 75 users rated 1,698 unique objects on a scale of 1 to 5 for usefulness; each of these objects was rated 1.3 times on average. Additionally, 79 users tagged 1,838 unique objects with 12,041 tags in total; consequently each object was assigned with 6.5 tags on average. Similarly to MACE, the learning objects can hold classification keywords from a controlled vocabulary additionally to the free text tags. 97.97% of the learning objects hold tags or classification, 95.53% hold tags, 69.04% hold classifications and 66.6% hold both. Additionally, 73 users rated 1,597 learning objects, each learning object was rated 1-10 times (on average 1.34 times) and each user rated 1-108 learning objects (on average 29.29) which results in a user-item rating matrix with a sparsity of 98-17%.

5 Experiment 1: Usage Context and Content Similarity

5.1 Methodology

Create Usage Context-based Similarities. We calculate the usage context-based similarity for all learning object pairs in the MACE and in the TravelWell dataset, respectively, as described in chapter 3. We start with calculating the co-occurrences and their significance values for each object using the association measures MI, log likelihood, and corrected χ^2 (see subchapter 3.2). We use two ways to select the most significant co-occurrences for each object. First, we vary the co-occurrence vector sizes from 1-150 for TravelWell and from 1-1000 for MACE. The vector sizes for MACE get bigger as for TravelWell since the MACE dataset comprises more learning objects. Second, for each object, we calculate an object-specific threshold by averaging the significance values of all its co-occurrences. We calculate one threshold for each object and not one threshold for all objects, because the significance values vary depending on the times an object was used. Thereafter, we calculate the usage context-based similarity of each object pair by comparing their co-occurrence vectors using the cosine similarity (see subchapter 3.3).

Create Semantic Metadata-Based Similarities. We calculate the semantic metadata-based similarity of all object pairs in MACE and TravelWell to get a reference value for evaluating the usage context-based similarities. We do so by taking the tags and classifications into account. Since an object cannot be tagged more than once with the same keyword, we create a binary vector for each object and use the Jaccard similarity [24] for calculating the pair-wise semantic metadata-based similarity.

Calculate the Correlation of Both Similarity Measures. We use the semantic metadata-based similarity as tentative "gold standard". Even though the semantic metadata-based similarity cannot be a perfect representation of the "real" similarity, e.g. because some learning objects hold only one or two tags, we assume it to be a good approximation. In order to prove our hypothesis that usage context similarity implies content similarity, we calculate the Pearson correlation coefficient [25] between the semantic metadata-based and the usage context-based similarity distribution, see formula 5 with X being the set containing all usage context-based similarities with $\bar{x}$ as mean value and Y being the set containing all metadata similarities with $\bar{y}$ as mean value.

$$r_{XY} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{5}$$

We only calculate the Pearson correlation coefficient for learning objects that hold at least one semantic metadata-based and one usage context-based similarity to another learning object (not necessarily the same learning object) that is

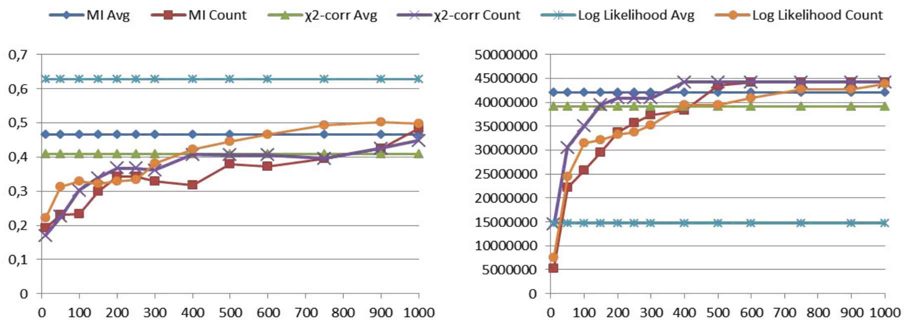


Fig. 1. MACE: Pearson correlation coefficient and amount of considered object pairs

greater than zero. Thus, we exclude learning objects that do not hold a) sufficient semantic metadata and/or b) a sufficient amount of significant co-occurrences to establish semantic metadata- and usage context-based relations, respectively, to other learning objects. Even if a learning object holds semantic metadata, it might not be able to create a semantic metadata-based similarity to another object greater than zero since 72.96% of the tags assigned to MACE objects and 73.87% of the tags assigned to TravelWell objects are unique and cannot be used to compare objects. The same holds true for the usage context-based similarity, the less objects a co-occurrence vector holds, the less is the chance to find similarities to other objects. However, when at least one similarity was found for a learning object, the object is used for the calculation of the Pearson correlation coefficient with all possible object pair combinations.

5.2 Results

MACE. Figure 1 shows a) the Pearson correlation coefficients for the semantic metadata-based and the usage context-based similarities that are calculated with different association measures (MI, log likelihood, and corrected χ^2) and varying vector sizes ("count" means that a fixed n was chosen, "avg" means that an object specific threshold, i.e. its average significance value, was used, see section 3.2) as well as b) the amount of object pairs that can be considered for each combination of association measure and vector size for the MACE dataset.

As could be assumed, the more co-occurrences are considered as significant and, thus, are used to describe an object, the more usage context-based similarities can be detected between object pairs. Additionally, the correlation with the semantic metadata-based similarity increases with the co-occurrence vector size, whereas from a certain vector size on (here: 1000), the amount of similar object pairs found and the correlation gets stable. This is due to the fact, that most objects hold less than 1000 co-occurrences and are already described to their full extend. With vector size 1000 (which means at maximum 1000) the real average vector size is 213 (log likelihood), 255 (MI), and 261 (corrected χ^2). Interestingly, at vector size 1000, log likelihood performs best in terms of correlation (0,4973),

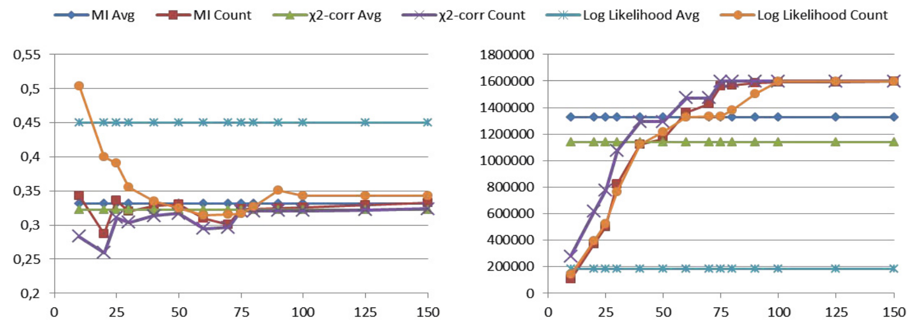


Fig. 2. TravelWell: Pearson correlation coefficient and amount of considered object pairs

followed by MI (0,4844) and corrected χ^2 (0,4487), which shows that "the more, the better" does not hold true, but the insignificant co-occurrences which can be considered as noise must be filtered to reach sufficient results.

The log-likelihood measure with the average threshold performs best in terms of correlation with a value of 0,629. However, it can only calculate similarity values for a subset of the available objects (14.707.176 object pairs). When comparing the association measures MI and corrected χ^2 with the average threshold, MI performs significantly better with a correlation value of 0,4653 (compared to 0,4079) and 42.067.378 object pairs (compared to 39.121.435).

TravelWell. Figure 2 shows a) the Pearson correlation coefficients for the different association measures and varying vector sizes as well as b) the amount of object pairs that can be considered for the calculation of the Pearson correlation coefficient for the TravelWell dataset. Similar to MACE, the amount of objects for which usage context-based similarities can be established with other learning objects increases with the vector size, but the results stabilize between vector size 100 and 150. This is due to the fact that the TravelWell dataset only holds 1.925 learning objects whereas the MACE dataset holds 12.176 objects.

Apart from the outliers when using very small vector sizes (10-25), the correlation coefficient increases with the vector size. For vector size 150, the association measure log-likelihood performs best with a correlation coefficient of 0,3432, followed by MI (0,3324) and χ^2 (0,3235). When using the average threshold, log-likelihood performs significantly better than the other measures in terms of correlation (0,4506) but only for a subset of objects (180.300 object pairs). MI outperforms χ^2 with a correlation coefficient of 0,3317 (compared to 0,3227) and 1.326.006 object pairs (compared to 1.140.805).

5.3 Interpretation

The resulting coefficients can be described as medium and as we regard a large sample of objects, the coefficients can be said to be representative although no tests for bivariate normality have been undertaken [26]. Moreover, because of the big sample size the correlations are significant on the 5%-level (p=.05).

A further important point is the fact that the considered metadata can only be interpreted as a shallow content representation. Thus, it is possible that the correlation between semantic metadata- and usage-based similarity represents a lower bound for the real correlation between content and usage context.

For the MACE dataset, the reached correlation coefficients are significantly higher than for the TravelWell dataset; we assume this is due to the fact that the MACE dataset holds more detailed usage data, e.g. each learning object access is tracked and not only the metadata provision activities rating and tagging as in TravelWell. Additionally, the MACE usage data offer precise timestamps for the collected events, whereas TravelWell only provides the date; thus, the sessions are more accurate in the MACE dataset.

The three association measures MI, log-likelihood and χ^2 behave quite similarly compared to each other for the MACE and the TravelWell dataset. For both datasets, the best performing set-up in terms of correlation coefficient and amount of object pairs is the log-likelihood measure in combination with a "large" vector size, whereat the meaning of "large" must be defined depending on the dataset. For MACE and TravelWell, a vector size of about 7% of the number of distinct objects in the dataset is recommended. This set-up is followed by the association measure MI in combination with the average threshold which has the advantage that no parameter for the vector size must be defined, thus, we use this set-up in the following experiment.

6 Experiment 2: Usage Context-Boosted Recommendations

6.1 Methodology

Usage Context and Content-Based Boosting. Melville et al. [13] introduced content-boosted collaborative filtering, i.e. a feature-augmented hybrid recommender system (see subsection 2.4) that uses the given rating history of users and content information of objects to predict the missing ratings in a user-item rating matrix using a content-based recommender. This enhanced matrix is then used as input for traditional collaborative filtering.

In the previous section we show that the usage context-based similarity gives an indication for the semantic similarity of learning objects pairs. Here, we use a usage context-based recommender to predict the missing ratings in the user-item rating matrix (see subsection 3.4). This matrix is then used as input for several recommendation approaches which are described in the next subsection. Thus, the usage context-based boosting is similar to the content-based boosting but does not require any content information which are often not given in a sufficient amount for learning objects.

Recommendation Approaches. In order to evaluate our approach we use the standard collaborative filtering methods item-based (with adjusted cosine similarity) and user-based (with Pearson correlation based similarity) collaborative filtering. Additionally, we tested the matrix factorization methods (MF) offered by the PREA toolkit [27] (i.e. Single Value Decomposition (SVD), Non-negative MF, Probabilistic MF, and Bayesian Probabilistic MF) as well as the MF methods offered by the Java port of the MyMediaLite Recommender System Library [28] (i.e. a standard MF as well as a Biased and a Factorized MF). Based on the performances of the different methods on our test sets and to not overload the diagrams, we choose to present the SVD [8] method from the PREA toolkit and the Biased Matrix Factorization (BMF) [9] from the MyMediaLite toolkit.

Experimental Set-up. We perform a 5-fold cross evaluation to compare the recommendation approaches with their usage context- and content-boosted versions for the MACE and the TravelWell dataset. In order to do so, we calculate the predictive evaluation metric Root Mean Squared Error (RMSE), see formula 6. Predictive accuracy metrics measure the deviation between a predicted rating $p(u,i)$ and the user's true rating $r(u,i)$ of item i for all users U and their ratings $P(u)$ in the test set. The most commonly used predictive accuracy metrics are Mean Absolute Error (MAE) and RMSE. The MAE measures the average absolute deviation of the expected ratings and the users' true ratings, whereas RMSE squares the deviations before they are averaged. Thus, the RMSE gives a higher weight to large errors.

$$\text{RMSE} = \sqrt{\frac{\sum_{u \in U} \sum_{i \in P(u)} (p(u,i) - r(u,i))^2}{\sum_{u \in U} |P(u)|}} \tag{6}$$

6.2 Results

MACE. Table 2 shows the results for the MACE dataset. Due to the sparsity of the original user-item rating matrix, the baseline algorithms IBCF, UBCF and SVD are only able to predict ratings for a small subset (14.8% - 24.9%) of the ratings in the test sets. BMF predictes a rating for each possible user-object combination. For all underlying algorithms, but the IBCF, the boosted approaches are able to decrease the RMSE up to 16.39% (usage context-based) and up to 11.66% (content-based), respectively. Additionally, up to 4.5 times as much ratings are predicted with the boosted approaches. For all approaches, the usage context-based boosting performs better than the content-based boosting in terms of RMSE and quantity of predicted ratings.

TravelWell. Table 3 shows the results for the TravelWell dataset. Similarly to the MACE dataset, the baseline algorithms, with BMF as exception, can only create predicted ratings for a subset of the ratings in the test sets whereas all boosted approaches create predicted ratings for about three times as many user- combinations. For the IBCF, SVD and BMF algorithms, both approaches are able to decrease the RMSE up to 2.81% (usage context-based) and up to 4.66% (content-based). In contrast to the MACE data set, the content-boosted approach outperforms the usage context-boosted approach for all algorithms.

Table 2. MACE: Comparison of the recommendation approaches

	Baseline		Usage Context-Boosted		Content-Boosted	
	RMSE	Quantity	RMSE	Quantity	RMSE	Quantity
IBCF	1.0090	14.8%	1.0562	67.9%	1.0675	63.9%
UBCF	1.0816	14.8%	1.0792	67.9%	1.1329	63.9%
SVD	1.2975	24.9%	1.0848	70.4%	1.1462	67.9%
BMF	1.1144	100%	1.0583	100%	1.0821	100%

Table 3. TravelWell: Comparison of the recommendation approaches

	Baseline		Usage Context-Boosted		Content-Boosted	
	RMSE	Quantity	RMSE	Quantity	RMSE	Quantity
IBCF	0.8909	31.5%	0.8769	94.7%	0.8494	93.3%
UBCF	0.8946	31.5%	0.9255	94.7%	0.8537	93.3%
SVD	0.9095	35.8%	0.8839	94.8%	0.8712	93.4%
BMF	0.8656	100%	0.8533	100%	0.8342	100%

6.3 Interpretation

The MACE portal offers more extensive usage data than the TravelWell portal, i.e. more tracked events including a timestamp. Thus, as shown in chapter 5, the MACE usage data is better suited to create usage context-based similarities that imply semantic similarities. In contrast, the objects in the TravelWell dataset hold more content information than the objects in the MACE dataset. As a result, all usage context-boosted approaches perform better than the content-boosted approaches for the MACE dataset, whereas for the TravelWell dataset, the content-boosted approaches perform better.

For both datasets, the MF methods SVD and BMF profit from the usage context-based and the content-based boosting approach. The standard collaborative filtering approaches IBCF and UBCF are more sensible concerning expected ratings in the user-item rating matrix that differ from the true user ratings.

To conclude, boosting can be recommended for the use in learning portals, especially in combination with the Biased Matrix Factorization (BMF) approach. If the collected usage data is fine-grained, i.e. all events concerning a learning object are stored and the timestamp is given and/or only sparse semantic metadata information is given for the learning objects, the use of usage context-based similarity can even outperform the use of semantic metadata-based similarity.

7 Conclusion

In this paper, we introduce a new way to calculate similarities between learning objects by considering their usage contexts. The usage context of a learning object is defined by the objects it significantly co-occurs with in user sessions. This way, learning objects are similar if they co-occur with the same learning objects. This also means, two learning objects can be similar even if they were never used together in the same session. Our hypothesis that usage context-based similarity is an indication for content similarity is supported by the experimental evaluation using the datasets MACE and TravelWell.

The usage context-based similarity can be used in several ways. Here, we boost existing recommendation approaches by filling up the underlying user-item rating matrix with expected ratings that are calculated using the usage context-based similarity. We evaluate this approach against the original recommendation approaches and their content-boosted equivalents. This evaluation shows, that the usage context-based similarity is able to even outperform the content-based similarity if the usage data is fine-grained enough. Thus, simple log files of systems providing learning resources are sufficient to calculate similarities between the resources that can be incrementally updated.

The results motivate us to further develop this approach. First, we will try to find a suitable way to combine usage context and content-based similarities to not waste any available information. Additionally, we plan to evaluate our approach in a running system to gather more insights about the usefulness of the created recommendations.

References

1. Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., Koper, R.: Recommender systems in technology enhanced learning. In: Rokach, L., Shapira, B., Kantor, P. (eds.) Recommender Systems Handbook, pp. 387–415. Springer (2010)
2. Verbert, K., Drachsler, H., N.M.: Dataset-driven research for improving tel recommender systems. In: 1st ACM International Conference on Learning Analytics and Knowledge, pp. 44–53 (2011)
3. Meteren, R.V., Someren, M.V.: Using content-based filtering for recommendation. In: Proceedings of MLnet/ECML2000 Workshop (2000)
4. Pazzani, M., Muramatsu, J., Billsus, D.: Syskill and webert: Identifying interesting web sites. In: Proceedings of Thirteenth National Conference on Artificial Intelligence, AAAI 1996, pp. 54–61 (1996)
5. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th International Conference on World Wide Web, pp. 285–295 (2001)
6. Linden, G., Smith, B., York, J.: Amazon.com recommendations - item-to-item collaborative filtering. IEEE Internet Computing 7(1), 76–80 (2003)
7. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer 42(8), 30–37 (2009)
8. Paterek, A.: Improving regularized singular value decomposition for collaborative filtering. In: Proceedings of KDD Cup and Workshop (2007)
9. Rendle, S., Schmidt-Thieme, L.: Online-updating regularized kernel matrix factorization models for large-scale recommender systems. In: Proceedings of the 2008 ACM Conference on Recommender Systems, RecSys 2008, pp. 251–258. ACM (2008)
10. Burke, R.: Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction 12(4), 331–370 (2002)

11. Mobasher, B., Jin, X., Zhou, Y.: Semantically enhanced collaborative filtering on the web. In: Berendt, B., Hotho, A., Mladenič, D., van Someren, M., Spiliopoulou, M., Stumme, G. (eds.) EWMF 2003. LNCS (LNAI), vol. 3209, pp. 57–76. Springer, Heidelberg (2004)
12. Tran, T., Cohen, R.: Hybrid recommender systems for electronic commerce. In: Knowledge-Based Electronic Markets, AAAI Workshop, pp. 78–83. AAAI Press (2000)
13. Melville, P., Mooney, R.J., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. In: Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI 2002), Edmonton, Alberta, pp. 187–192 (2002)
14. de Saussure, F.: Course in General Linguistics. Open Court Publishing, Chicago (re-print 1986)
15. Heyer, G., Quasthoff, U., Wittig, T.: Text Mining: Wissensrohstoff Text. Konzepte, Algorithmen, Ergebnisse. W3L-Verlag, Herdecke (2006)
16. Manning, C.D., Raghavan, P., Schütze, H.: An Introduction to Information Retrieval. Cambridge University Press (2008)
17. Evert, S.: The Statistics of Word Cooccurrences: Word Pairs and Collocations (Dissertation). University Stuttgart (2004)
18. DeGroot, M.H., Schervish, M.J.: Probability and Statistics, 3rd edn. Addison-Wesley, Boston (2002)
19. Dunning, T.E.: Accurate methods for the statistics of surprise and coincidence. Computational Linguistics 19(1), 61–74 (1993)
20. Evert, S.: Corpora and collocations. In: Corpus Linguistics. An International Handbook, vol. 2, pp. 223–233. De Gruyter, Berlin (2009)
21. Stefaner, M., Dalla Vecchia, E., Condotta, M., Wolpers, M., Specht, M., Apelt, S., Duval, E.: MACE – enriching architectural learning objects for experience multiplication. In: Duval, E., Klamma, R., Wolpers, M. (eds.) EC-TEL 2007. LNCS, vol. 4753, pp. 322–336. Springer, Heidelberg (2007)
22. Schmitz, H.C., Wolpers, M., Kirschenmann, U., Niemann, K.: Contextualized attention metadata. In: Roda, C. (ed.) Human Attention in Digital Environments. Cambridge University Press (2011)
23. Vuorikari, R., Massart, D.: Datatel challenge: European schoolnet's travel well dataset. In: 1st Workshop on Recommender Systems for Technology Enhanced Learning, RecSysTEL (2010)
24. Salton, G., McGill, M.J.: Introduction to modern information retrieval. McGraw-Hill, New York (1983)
25. Pearson, K.: On further methods of determining correlation. Cambridge University Press (1907)
26. Bortz, J.: Statistics. Springer, Heidelberg (1993)
27. Lee, J., Sun, M., Lebanon, G.: A Comparative Study of Collaborative Filtering Algorithms. ArXiv e-prints (May 2012)
28. Gantner, Z., Rendle, S., Freudenthaler, C., Schmidt-Thieme, L.: MyMediaLite: A free recommender system library. In: Proceedings of the 5th ACM Conference on Recommender Systems, RecSys (2011)

Design Principles for Competence Management in Curriculum Development

Kai Pata, Kairit Tammets, Mart Laanpere, and Vladimir Tomberg

Tallinn University, Institute of Informatics, Narva road 25,
10120, Tallinn, Estonia
{kpata,Kairit.Tammets,martl,vtomberg}@tlu.ee

Abstract. This paper discusses the ontology-based competence management in the context of curriculum development. The data was collected through two participatory design experiments, which were conducted to develop and formatively evaluate the designs and scenarios for computer-supported curriculum development in higher education context. These experiments involved the use of prototypes of software tools developed in two different research projects: Intelleo and ePoAbi. As a result, the paper proposes ontology-based design principles to computer-supported curriculum development, course development and curriculum evaluation.

Keywords: curriculum development support, competence management.

1 Introduction

Curriculum design and management is one of the most challenging and least automated areas in education where semantic technologies seem to have unrealized potential. Several authors have investigated the possibilities of semantic technologies in curriculum design, management and evaluation [1-7]. These studies propose ontologies to be used in the curriculum development systems for mapping courses and resources from the curriculum, managing learning outcomes, and modeling curriculum management. However, this paper deals with the gap of ontology-based harmonization of top-to-down and bottom-up approaches in curriculum development and management.

Our study is based on two participatory design experiments involving two original software tool prototypes. Organizational Policy Tool (OPT) was developed within FP7 funded international R&D project IntelLEO (http://intelleo.eu) to support mapping and evaluation on competence standards in extended learning organizations (e.g. temporary partnership of business enterprise and university). ePoAbi tool was developed within a national-level research project ESF Primus to support Web-based composition of course syllabi in accordance with competence standard requirements. As a result of this study, we propose the scenarios for curriculum and course development and evaluation, the curriculum ontology and the design principles for competency-based curriculum design tools.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 260–273, 2013.

2 Competence Management in Curriculum Development

2.1 Curriculum Development Concepts and Process

In June 1999, representatives of the Ministers of Education of 29 European countries convened in Bologna, Italy, to formulate the Bologna Declaration, which aimed at establishing a common European Higher Education Area (EHEA) [8]. The Bologna process spells out a number of ‘action lines’ in which learning outcomes play an important role [9]. One of the decisions was that education in higher education institutions throughout the EHEA should be based on the concept of learning outcomes, and that curricula should be redesigned to reflect this. These outcomes would become the foundation for decisions about the curriculum, instruction, assessment, staff development and so on. Nationally OBE has been implemented through the requirements in the National Higher Education Standards. These standards are obligatory when designing curricula and courses in higher education institutions, and are used in the curriculum evaluation process. Van der Horst and McDonald [10] noted that one of the problems of OBE lies in the teachers’ ability to translate the imprecisely worded outcomes into their teaching and learning practice.

Curriculum is made up of a number of parts, what in general fall under the broad headings of content, learning experiences, objectives and assessment [11]. Broader concept of curriculum defines it as a plan for learning [12] that specifies how learning takes place, considers central rationale, the aims and objectives, content, organization, and evaluation of learning [13]. It contains a sophisticated blend of educational strategies, course content, learning outcomes, educational experiences, assessment, the educational environment and the individual students’ learning style, personal timetable and the program of work [14]. Course is a subset of a program of study (equivalent to a module or unit of study) within a curriculum. Typically, course design starts with the definitions of the competences that must be met at the end of the course.

Curriculum is not something static, and permanent change is one of its key features. Lu [6] has detected that it is not easy to keep track of the curriculum evolution – the information of the curriculum design may get lost when interpreted by different stakeholders. Often the curriculum design is hindered by the lack of shared vocabulary among different stakeholders (such as curriculum board and the teachers who prepare the courses) to communicate about the curriculum.

Figure 1 below illustrates the core part of the curriculum ontology in the form of a concept map, derived from the literature review and adapted to Estonian higher education standard.

Developing a university curriculum and keeping it up to date is guided by two counter-directional factors:

I. The stabilizing factor: Matching the curriculum goals and the course goals to the existing standards in certain domains

Informed and intentional curriculum design is guided by a vision, what considers university and national strategic plans, policies, priorities and quality processes,

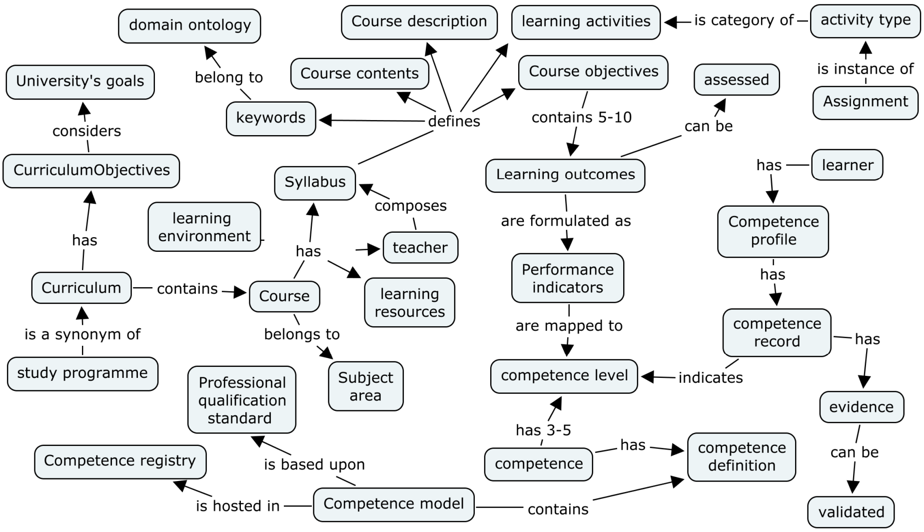


Fig. 1. Curriculum ontology in the form of a concept map

external accreditation processes and domain knowledge and competences. The head of the curriculum and the curriculum board should define overall goals and aims for the curriculum, stating specific measurable knowledge, skill/ performance, attitude, and process objectives for it, and defining how these goals might be achieved in different course modules. Learning goals and outputs for every course and subject should be compatible with those criteria. Learning outcome is an expression of what a student will demonstrate on the successful completion of a course. Learning outcomes are related to the level of the learning; indicate the intended gain in knowledge and skills that a typical student will achieve and should be capable of being assessed. Learning outcomes are more about the learning that is actually to be achieved by the learner. Outcomes ten formulated as competences. Sampson and Fytros [15] define competences as personal characteristics (e.g. skills, knowledge, attitudes) that an individual possesses or needs to acquire, in order to perform an activity within a specific context, whereas performance may range from the basic level of proficiency to the highest levels of excellence.

II. The change factor: Advancing the existing competence conceptualizations and creating new competences.

Boud [16] points out that it's the courses not a narrow top-to-down operational competency-based approach that is guided only by curriculum learning outcomes that gives an educational approach to the curriculum. Furedi [17] argues that curriculum objectives should be flexible enough to allow 'individual scholars to pursue their passionate interests' [18, p. 2] within the curriculum objectives, however, very few course developers examine 2the full curriculum document.

Hsiao et al., [19] have suggested three stages to develop a practical competency oriented curriculum: the first is the planning stage; the second is designing stage and the third stage is to choose the representative subjects in this program, to develop the teaching materials and to proceed with the experimental instructions. In the first stage the board discusses and analyzes, which courses should be taught and what competences required for labor market. In the next change the development and design of the curriculum takes place. The national and international competence standards in this domain area serve as the main guiding factors in the goal development, enabling standardization of learning goals, and the further evaluation of curricula, based on these standards.

If the curriculum board would be willing to bring in the change factor into the curriculum development, the course developers need to be provided with the freedom creating new learning outcomes for course programs and proposing these to be integrated to the curriculum learning outcomes. Thus, after the curriculum goals have been defined at competence level, the initiative should be given to the course developers who must take responsibility for developing the particular course goals. Roche [20] have also indicated that collaboration between the curriculum designers and academics in curriculum development process is an efficient strategy that promotes creating curriculum ownership [2]. Assuming that each course developer is an expert in this domain, and is involved into the international research and development besides teaching, would guarantee that they could embed this front-line knowledge and expertise into their courses. It is highly unlikely that the existing competence descriptions, derived from the normative standards into the curriculum would cover all the required competences needed at the courses. The course developers must define new learning goals for their courses that do not appear in competence standards, and backward matching to the curriculum and standard must be conducted. Zervas and Sampson [21] have described that both the available and required competences for a certain role, job or function should be kept updated in formal standards. In the curriculum development process, it means that matching new competences, searching, mapping and filtering them, is needed.

2.2 Managing Competences in Outcome-Based Education

Several studies proposed systems that support developing curricula using the competence-management approach. Ronchetti and Santi [1] developed the competence-ontology based strategy for managing, inspecting and monitoring a full course of study (a program). They proposed to use ontologies as a management tool for coverage of a full study course, one that would be useful in curriculum improvement and program reviews. In their experiment an analysis of the syllabi of the computer science bachelor courses in Trento University was conducted, associating the topics in the curriculum courses with the reference ontology. They indicate two possible objections to using ontologies in curriculum development: a) the reference ontologies take considerable effort and time to develop, and since curricula need to evolve, the ontology might not support new topics that emerge, b) creating an ontology and indexing all the courses' content against it requires much effort. Paquette [3] proposed the ontology and software framework TELOS for competency modeling and resource management to support learning program development as an assembly-and-coordination system, which uses competencies, their groupings to units (called courses), and the estimated competency gap to define the structure of the new program. His ontology for competency modeling defines competencies as statements that someone, and more generally some resource, can demonstrate the application of a generic skill to some knowledge, with a certain degree of performance. The knowledge/competency structure presents a hierarchy of concepts in an application domain, with their associated entry and target competencies. Lu [6] developed the ontology-aware course-consulting system On2C for curriculum evolution that uses the conceptual model of the evolutionary processes of the curriculum. Barrera et al. [5] described the semantic system DIOGENE, which aims at supporting the design of competency-based curricula. Vaquero et al. [4] proposed a methodology intended to improve course design using semantics and case-base reasoning techniques. Dexter and Davies [2] developed the ontology-based curriculum knowledgebase - the CRAMPON system – for managing complexity and change in curriculum management in medical studies. The knowledgebase system manages and maintains the complex interrelationships between curriculum content, dynamically updated medical cases for enquiry based learning and their intended learning outcomes, and assessment based on the outcomes. Any competency may be ‘tagged’ with values such as curriculum aims, outcome etc., which allows detecting ‘gaps’ and ‘overcrowding’ in the curriculum. Dexter and Davies [2] indicate that both academic staff and the students might benefit from such a system in preparation for quality reviews, and planning and executing changes to the curriculum; and making appropriate aspects of the curriculum intuitively visible to facilitate self-directed learning. For instance, they describe the mapping issues between the standard and the curriculum as one of the challenges - updating any aspect of the curriculum in line with new standard guidelines involves identifying and tracking all the related parts in course programs on which that aspect may impinge and agreeing the changes within the curriculum development team. CRAMPON system enables the following: defining a core curriculum; mapping of external criteria onto the curriculum; linking assessments to the curriculum; mapping cases to the curriculum; allowing more systematic, rigorous and transparent curriculum content design and revision and search; extending ‘ownership’ of the curriculum to a larger group of teachers who would otherwise never be consulted; support widespread collaboration and consultation on curriculum expertise inside and outside the institution, allow managing and quality assessment of the curriculum. From the learner’s side the CRAMPON system has a connection to their e-portfolios and allows browsing curriculum for the identification of strands of activity in the curriculum and enables them to plan learning. Karunananda, et al. [7] described the OntoCD as an ontology driven approach for curriculum development that runs on protégé ontology development toolkit. The expected users of OntoCD are curriculum designers, persons who revise curriculum and maintain degree programs. The features of the system are loading the benchmark ontology for a curriculum, loading skeleton curriculum or to start from scratch, validating skeleton curriculum against ontology (missing knowledge areas, core/electives, credits), creating modules from knowledge area and distributing them across the study, annotating modules, evaluating curriculum.

Three aspects may be summarized based on these studies:

The targets for Curriculum development tools are mainly academic staff and evaluators of the curricula, the tools are meant for promoting taking ownership of the curriculum based on competence management, and allow the two-way curriculum development, evolution and evaluation. However, some studies [2] suggest that the users for new type of curriculum development tools may target also self-directed learners helping to search across the curriculum and providing them with the enacted curriculum output in their portfolios.

Most of the tools focus on curriculum management from its narrower perspective (managing competences). Few systems support also focusing on curriculum content and learning experiences, besides objectives and assessment and conceptualize learning program development as an assembly-and-coordination system, using competencies to define the structure, learning cases and the resources of the program [1-3]. Some systems target rather the course design [4].

The ontology support is provided for the following purposes: reference ontologies catch the central concepts and relations of a domain and have been used for mapping courses and resources from the curriculum [1-2]; competence ontologies, have been used for managing learning outcomes [3-5]; benchmark ontologies for the curriculum have been used to model curriculum management aspects and phases [6-7].

3 Methods

We evaluated the curriculum development tool prototypes OPT and ePoAbi formatively using the case of Educational technology master curriculum in Tallinn University, Estonia. OPT was used for defining the competence lists of an Educational technology master curriculum; for harmonizing curriculum competences with competence lists provided by other organizations, such as the ISTE ICT competence standard for teacher for accreditation purposes, and; proposing competences to the standard. ePoAbi was tested for harmonizing the competences in a certain course with the acknowledged competence standards; harmonizing the competences, learning activities and assessment types in a certain course.

OPT evaluation was done as following: In the design experiment the head of the curriculum tested the original competence-harmonization tool OPT for harmonizing the Educational Technology master curriculum with the International ISTE ICT competence standard for teachers. Secondly, a competence standard expert from the Tiger Leap foundation, responsible for Estonian teachers' ICT competences, evaluated the master curriculum based on the standard. Interviews with them were conducted. ePoAbi prototype was evaluated with 7 course developers of the Educational Technology curriculum. The sample was created from the staff of the compulsory subject domain courses in this curriculum. Initially the system was populated with some example learning-outcome, -activity and assessment descriptions that were collected as part of the qualitative study among the elearning trainers [18]. Each user revised their course program – the learning objectives, -activities and assessment types using the ePoAbi system. This enabled them experiencing the course-management system in real context. Each course developer could populate the system with the learning outcome, -activity and assessment descriptions and benefit from those added by other colleagues. We conducted the post-questionnaire with the ePoAbi users.

4 Evaluation of the Curriculum Development Scenarios

4.1 The Evaluation of the OPT Tool

I. Curriculum development scenario has the following goals:

a) Based oon standards, defining the competence lists for the curriculum
b) Proposing curriculum competences to the course programs
c) Harmonizing the competences in the curriculum with the competences formulated in the course programs of this curriculum
d) Proposing emerging competences from courses to the competence standard
e) Harmonizing the course competences, learning activities and assessment types across the curriculum

In the initial phase of curriculum development the competences and competence descriptions should be acquired from the different competence standards in order to compile the learning goals for the curriculum (see figure 2). Our experiment showed that this full mapping between two standards might be a very time-consuming activity.

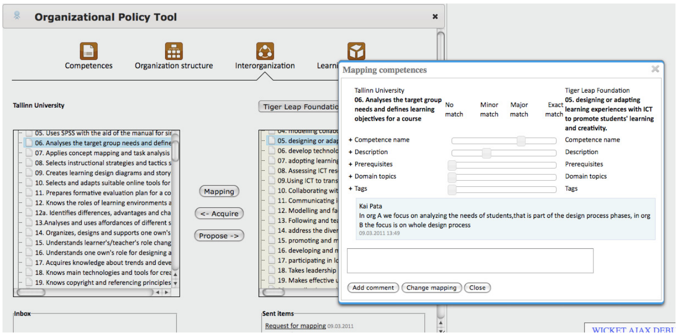


Fig. 2. Mapping the curriculum competences on the ISTE ICT competence standard for teachers

Since the list of competences in the standard/curriculum is decided by the board/commission, the direct uptake of one competence to the other organization's competence list is impossible. Rather the proposed competences from standard should be stored in the *temporary store* and considered while a) changing the competence descriptions or b) acquiring the whole competence description. This process may be accompanied by the further discussions between the competence managers in two organizations. The curriculum competence manager and the competence standard manager may take a partial responsibility for the competence descriptions in another organization. For example, they may see that some emerging competences that are present in their learning goals list are missing in the standard or that the curriculum does not have learning goals associated to some competences in the standard. They may also see the need to rephrase some competence descriptions of another organization. In this case they may propose that the other organization could acquire some competence descriptions. The explanation, why this competence may be needed should be added to the proposal. We found that long discussion threads between the competence managers do not appear. Rather, one comment and short response are common. The users also mentioned that it is important to keep these comments as notes associated with the competences for the future purposes. Especially in case of mapping the competences the reason for mapping must be stored for the future.

These compiled competences form the curriculum goals may be made available as templates for the lecturers/teachers who start planning the course goals (see the Course development scenario for ePoAbi below). It would be very useful to see the emergent hierarchy of competences of the curriculum sorting them by modules or by domain concepts. Curriculum competences should be tagged using the domain ontology, as well as with the curriculum module hierarchies. Tags will enable filtering the set of competences that belong into one module, or to certain domain concept that supports curriculum evaluation activities.

II. Curriculum evaluation scenario has the following goals

f) Estimating the match between curriculum competences with competence standards, such as the ISTE ICT competence standard for accreditation purposes
g) Estimating the balance of competences, -learning activities and assessment types at the curriculum

The curriculum evaluation scenario requires analyzing to what extent the curriculum and its modules (the competences in the curriculum) cover certain normative competences in the standard. In evaluation, matching competences in the curriculum and the standard may stay uni-directional and serves for the purpose of analyzing the match and overlap between curriculum and standard competences. The evaluation would become more useful, if the overall match between the sets of competences (for example by one competence in the standard, by curriculum modules or by the competence groups) could be analyzed and visualized.

Secondly it is important to evaluate how the learning objectives of the curriculum are achieved and evaluated. For this purpose, getting an analytical overview of the competences, -activities and assessment types used at different courses for seeing which are more frequent or less used across the curriculum is needed. In ePoAbi tools (presented below) we tested the functionalities of viewing from analytics, which combinations of competences, -activities and assessment types have been frequently used together in the curriculum, which are more used competence, -activity and assessment type formulations, and which formulations have been viewed frequently by other lecturers in the curriculum.

4.2 The Evaluation of the ePoAbi Tool

III. Course development scenario has the following goals

h) Harmonizing the competences in a certain course with the acknowledged competence standards and the curriculum competences;

i) Harmonizing the competences, learning activities and assessment types in a certain course, matching competences with (one or several) learning activities and assessment types

ePoAbi supports course developers in formulating the competences according to the standard requirements and enables associating the learning activities and assessment types with the formulated competences. The course developer can initiate the course program from the scratch or use the previous program text and then formulate competences based on learner's expected outcomes in this course. System allows matching competences with one or several competence standard formulations or curriculum competences. In that phase course developer may view the curriculum board proposals to add or eliminate some competences from the course program and to get advice for (re)formulating competences from stored competence descriptions added by other lecturers from the same curriculum.

In the next ePoAbi supports the course developer in formulating the learning activities and assessment type for achieving of each competences. One or several learning activities or assessment types can be associated with the one competence of the course. There is a chance to get advice from the system for (re)formulating the learning activities and assessment types based on descriptions suggested by the system, which have been stored by other lecturers from the same curriculum. That can inspire course developers to take better ownership in the curriculum development and also use the mutual vocabulary with other course developers in this curriculum. After formulating and associating the competences, learning activities and assessment types the course developer may categorize each competence, activity and assessment type with the available competence ontologies and tag them with the suitable tags.

Analytics functionality allows course developers to view, which are more popular competences, learning activity and assessment type formulations and which formulations other lecturers have viewed frequently. The same functionality enables to see which combinations of competences, learning activities and assessment types have been frequently used together by other lecturers of the curriculum. For supporting the curriculum development process in general the course developer may propose new competences to be added to the curriculum and standard that are missing from there (see scenario I). In the end, the final course program with matched competences, learning activities and assessment types may be imported as the course program in .pdf format and delivered to the administrative unit.

We conducted the post-questionnaire with the ePoAbi users. The space limits of this paper do not allow presenting fully the data from this survey. However the main findings were the following: The course developers created learning outcomes, activities and assessment types for their courses but seldom viewed the social support which enabled to see the descriptions of other users. We distinguished two types of users – the critical but autonomous users considered their course programs sufficiently good and did not see the need for harmonization with standard learning outcomes; the supportive users were more open to radically change the learning outcome, -activity and assessment formulations of their course programs – this may be caused by the lower level of harmony between the learning outcomes, -activities and assessment types in their initial course programs. Both critical and supportive course designers considered enabling the bottom-up development and evaluation of the curriculum as the biggest merit of the system. The users suggested that social support functionality, collecting data about the curriculum as part of the course development process and the analytical functionalities should be provided when the curriculum is developed or annually evaluated internally.

5 Design Principles in Curriculum Design

5.1 Design Principles for Competence Management in Curriculum Design

Based on the literature review about curriculum development tools as well as our own experiments with the curriculum development tool prototypes OPT and ePoAbi we can propose the design principles what should be considered in developing tools for ontology-based harmonization of top-to-down and bottom-up approaches in curriculum development and management:

I. Translate learning outcomes of the curricula and courses to competences for managing curricula (see Fig. 1 for Curriculum Ontology).

II. Provide a shared vocabulary among different stakeholders (such as curriculum board and the teachers who prepare the courses) to communicate about the curriculum.

This may be supported by up-to down and bottom-up tagging competences, -learning activities and assessment types in curriculum development and course development phases. Secondly, we have proposed sharing the competence, -learning activity and assessment formulations and proposing these formulations to the course developers during the course development process, which might enable the development of shared vocabulary of this curriculum and extending 'ownership' of the curriculum to a larger group of teachers who would otherwise never be consulted because they do not belong to the curriculum board [2].

III. Manage the stabilizing factor in curriculum development.

The competence matching and harmonization tools must be provided at two directions – matching the curriculum with the competence standards and matching the courses with the curriculum. However, we found that matching at both levels is a time-consuming activity. Secondly, matching the course competences, learning activities and assessment must be supported. Matching tools need a temporal store. Challenges: There is no one-to-one match between competences, learning activities and assessment types: Some course developers use more than one learning activity to achieve a certain competence, others use the same learning activity to develop several competences and assess the same learning activity from each competence perspective.

IV. Manage the change in curriculum development for supporting the evolution of the curriculum

Provide tools to keep track of the curriculum evolution [6]. Updating any aspect of the curriculum in line with new standard guidelines involves identifying and tracking all the related parts in course programs on which that aspect may impinge and agreeing the changes within the curriculum development team [2]. Enable the course developers to take responsibility for developing the particular course goals [16] by providing flexible enough curriculum objectives that allow individual scholars to pursue their passionate interests' in the course development [17]. Collect from courses the competence conceptualizations and new competences that were developed by course developers, allow freedom in defining new competences, learning activities and assessment types for their courses that do not appear in curriculum standard, and enable backward matching to the curriculum and the standard.

V. Provide analytical tools to evaluate the curriculum internally and externally

Curriculum evaluation tools should enable searching, mapping and filtering information about competences, learning activities and assessment across curricula and its modules. For evaluation, match between the curriculum and its modules and courses to the standard competences as well as the coverage of certain competences at different courses and the match between the course competences, learning activities and assessment types must be made measurable. Analytical tools for curriculum evaluation should enable developing a quick overview of the state of the curriculum, and allowing a more detailed look for possible problem areas [1], for example visualizing the curriculum match with the standard and competence coverage with courses. For analytical tools any competence, learning activity and assessment type may be 'tagged' with values such as learning outcomes, modules as well as using the course developers' vocabulary [2].

VI. Match curriculum domain knowledge an the resources with the competences

Create and annotate modules from knowledge area and distribute them across the curriculum courses [7]. The knowledge/competency structure can present a hierarchy of concepts in an application domain, with their associated entry and target competencies [3]. Model and represent competence-related information in a machine-readable way so that it allows inter-exchange in a standard way between different system implementation [15]. The knowledgebase system may manage and maintain the complex interrelationships between curriculum content, dynamically updated learning cases and their intended learning outcomes, and assessment based on the outcomes, mapping cases to the curriculum [2]. Challenges: The reference ontologies take considerable effort and time to develop, and since curricula need to evolve, the ontology might not support new topics that emerge [1].

VII. Provide the co-development possibility for courses for creating synergy in curriculum evolution.

The same course design differs from one designer to other due to teacher's different viewpoints, and the re-use of knowledge and prior user experiences coming from different experts is rarely considered in course design [4]. Enable course developers to cooperate across courses for achieving some learning goals jointly. Enable the co-development of courses among several staff members responsible for the same course.

VIII. Make appropriate aspects of the curriculum intuitively visible to facilitate self-directed learning [2].

5.2 Curriculum Design Process

We outline the simplified Curriculum design process between different stakeholders: Curriculum head, Curriculum Board, Course Designer and the External Evaluator. The semantic support with competence ontologies is provided for harmonizing and matching competences of the Standard, the Curriculum and the Syllabi of the Courses and for the Curriculum Analytics. The Domain ontology is used to provide semantic support on relating learning resources and cases with the competences taught at the courses.

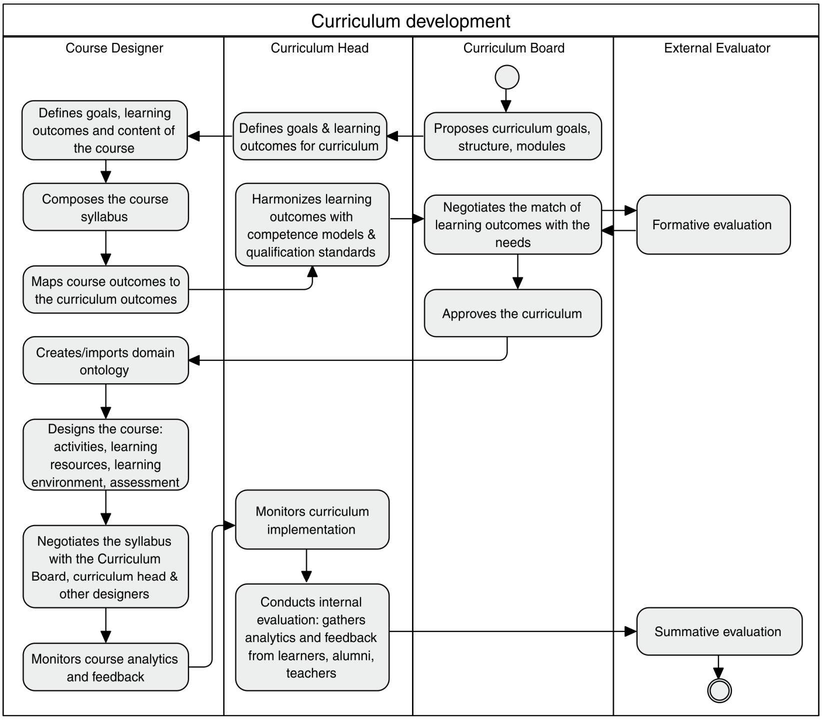


Fig. 3. Curriculum development process

6 Conclusion

In our paper we evaluated OPT and ePoAbi tool prototypes that use ontology-based approach to computer-supported curriculum development, course development and curriculum evaluation. Based on literature and the findings from formative evaluation we propose important design elements and the processes of competence based curriculum management. Our further research is directed towards integrating the two prototypes in order to provide support for curriculum development, management and evaluation as well as contribute to the evolution of competence standards. The future application areas for curriculum development tools might be supporting self-directed learners in choosing courses from the curriculum based on personal competence-plans. Another application area would be supporting course developers with the pre-composed learning environment configurations that meet to their course syllabi. We have made some developments towards this direction in the Dippler environment, where course environments can be composed by mutual involvement of Course developers who define learning outcomes, competences, activities and assessment formats and settings, while learners are enabled to integrate their personal social software tools to the course setting.

Acknowledgements. This study was funded by the Estonian MER grant 0130159s08, IST FP 7 project IntelLEO and ESF project Primus.

References

1. Ronchetti, M., Santi, J.: Curriculum Management and Review: An ontology-based solution.Technical Report # DIT-07-021 (2007)
2. Dexter, H., Davies, I.: An ontology-based curriculum knowledgebase for managing complexity and change. In: Ninth IEEE International Conference on Advanced Learning Technologies (2009)
3. Paquette, G.: An Ontology and a Software Framework for Competency Modeling and Management. Educational Technology & Society 10(3), 1–21 (2007)
4. Vaquero, J., Toro, C., Martín, J., Aregita, A.: Semantic Enhancement of the Course Curriculum Design Process. In: Velásquez, J.D., Ríos, S.A., Howlett, R.J., Jain, L.C. (eds.) KES 2009, Part I. LNCS (LNAI), vol. 5711, pp. 269–276. Springer, Heidelberg (2009)
5. Barrera, M., Montaño, N., Ramos, E.: An Ontological Approach to Support Design competency-based curriculum. In: Informatica (CLEI), 2012 XXXVIII Conferencia Latinoamericana. En, October 1-5, pp. 1–6 (2012)
6. Lu, W.: Ontology-aware Course Management for Curriculum Evolution Process in Higher Education. VERITE (March 13, 2009)
7. Karunananda, A.S., Rajakaruna, G.M., Jayalal, S.: Using Ontology for curriculum design and revision. In: Advances in ICT for Emerging Regions, ICTer 2012, pp. 137–144 (2012)
8. Kennedy, D., Hyland, A., Ryan, N.: Writing and Using Learning Outcomes: A Practical Guide. In: Bologna Handbook C 3.4-1 (2006), `http://www.bologna.msmt.cz/files/learning-outcomes.pdf`
9. Adam, S.: An Introduction to learning outcomes. In: Froment, E., Kohler, J., Purser, L. (eds.) EUA Bologna Handbook. Raabe, Berlin (2006)
10. van der Horst, H., McDonald, R.: Outcomes Based Education: A teacher's manual. Kagisto Education Publishers, Cape Town (1999)
11. Child, D.: Psychology and the Teacher, 7th edn. Continuum, Suffolk UK (2004)
12. Taba, H.: Curriculum development: Theory and practice. Harcourt, Brace & World, New York (1962)
13. Walker, D.F.: Fundamentals of curriculum: Passion and professionalism. Lawrence Erlbaum Associates, Mahjah (2003)
14. Harden, R.M.: AMEE Guide No. 21: Curriculum mapping: A tool for transparent and authentic teaching and learning. Medical Teacher 23(2), 123–137 (2001)

15. Sampson, D., Fytros, D.: Competence Models in Technology-enhanced Competence-based Learning. In: Adelsberger, H.H., Kinshuk, Pawlowski, J.M., Sampson, D. (eds.) International Handbook on Information Technologies for Education and Training, 2nd edn. Springer (2008)
16. Boud, D.: Creating a work-based curriculum. In: Boud, D., Solomon, N. (eds.) Work-based Learning. A New Higher Education?. SRHE and Open University Press, Buckingham (2003)
17. Furedi, F.: Where Have all the Intellectuals Gone: Confronting 21st Century Philistinism. Continuum International Publishing, UK (2004)
18. Tammets, K., Pata, K.: The trends and problems of planning outcome-based courses in elearning. In: Saar, E., Mõttus, R., Bern, E. (eds.) Higher Education at a Crossroad: The Case of Estonia, pp. 281–301. Peter Lang Publishers House (2013)
19. Hsiao, H.C., Chen, C.S., Huang, K.K.: Practice based curriculum development for technical institute. In: School International Conference on Information Technology Based Higher Education and Training, Kumamoto, Japan, pp. 451–457 (2001)
20. Roche, V.: Professional development models and transformative change: A case study of indicators of effective practice in higher education. International Journal for Academic Development 6(2), 120–129 (2001)
21. Zervas, P., Sampson, D.: The TENCompetence Observatory: An Enabling Technology for Common Description of Competences. In: Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies, Niigata, Japan (2007)

An Investigation of Two Methods for the Technological Mediation of Collaborative Peer Feedback in Higher Education

Andria Pavlou and Eleni A. Kyza

Cyprus University of Technology, Limassol, Cyprus
antriap89@gmail.com, Eleni.Kyza@cut.ac.cy

Abstract. Even though several studies have examined peer feedback through synchronous and asynchronous communication within the context of distance learning, few have addressed its contribution in higher education face-to-face settings. This proposal addresses this gap by examining peer feedback mediated through synchronous and asynchronous communication in the context of a face-to-face university course. Thirty-five undergraduate students participated in this study to provide feedback to each other's work through synchronous (chat) or asynchronous collaboration (in-context commenting). A variety of data were collected and included pre- and post-course surveys, videos and student interviews. The analysis of the results showed that students take advantage of the feedback they receive from their peers. As findings showed, peer feedback through asynchronous communication helped students learn, while peer feedback, mediated through synchronous tools, was important for interpersonal communication. Even when the feedback was not used to improve course work, students took it into serious consideration.

Keywords: Peer feedback, synchronous communication, asynchronous communication.

1 Introduction

Feedback is regarded as one of the most important contributors to learning. Hattie and Timperley [1] defined feedback as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (p. 81). A plethora of studies conducted so far confirm the significance of feedback during the teaching and learning process [2, 3].

As our current understanding of learning suggests that learning should be an active, student-centered process, peer feedback, as a method for facilitating reflection on one's practice is increasingly becoming more accepted [4]. In recent years, this idea found fertile ground among the various types of technologically-mediated communication. Several studies have investigated the issue of peer feedback through synchronous and asynchronous modes of communication [5, 6, 7]; however, most of them focused on studying the construct in distance learning settings. The present study addresses this issue by examining the topic of peer feedback provided by students enrolled in a face-to-face university level course.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 274–287, 2013.

2 Feedback in Education

The importance of feedback in learning has received considerable attention in education [8]. Matsangouras [9] emphasized the two-fold value of feedback. On the one hand, it provides information and guidance to the teacher. On the other hand, it helps students develop skills of communication, observation and self-evaluation [2].

Matsangouras [9] reports that an important condition for feedback is to state the targets clearly during lesson planning. Gibbs and Simpson [10] state that effective feedback should be timely so that its results can be exploited for further learning. Additional factors that may affect feedback are: (a) the type of feedback—oral, visual and kinesthetic [11]—(b) the quality and frequency of feedback [12] and (c) the source of the feedback—teacher or peer.

The source of feedback is probably the most determining factor in how students perceive feedback. The feedback from the instructor has been established as the primary source of assessment of students' work. Cho and MacArthur [12] report that the advantage of teachers' assessment is that, due to their experience, they give useful comments which may lead to the improvement of students' work. On the other hand, teacher feedback can sometimes be ineffective, as teachers may underestimate or overestimate the difficulty of a task assigned to the students, or overestimate the capabilities of students [12].

3 Peer Feedback

"While Vygotsky's zone of proximal development suggests that learning takes place in interaction with "more knowledgeable or more experienced individuals" (1978:86), later authors have come to widen this concept of expert knowledge to include peers with approximately the same level of language knowledge" [13] (p. 185).

Students as providers or receivers of feedback have more opportunities to learn, socialize and develop their cognitive skills by engaging in feedback processes [14]. Feedback from peers can support the learning process by providing an intermediate control of performance, indicating strengths and weaknesses and giving tips for improvement [15]. The benefits of peer feedback are not restricted to learning, but, they include the social aspect too. In peer-feedback contexts students create a small learning community [14], by engaging in discussions that make them feel more comfortable than in similar situations with the teacher. On the other hand, students' evaluation is often considered as less valid, a fact that may lead students to refuse to take into account the feedback from peers, especially when this is negative [16].

Several studies examined peer feedback and reached significant but, sometimes, contradictory results. For example, Liu and Carless [17] examined the contribution of peer feedback to Hong Kong students' learning using questionnaires and interviews. Results showed that a significant number of academics and students believed that students rarely evaluate each other in a valid way. Additionally, Xie, Ke and Sharma [18] found out that peer feedback in blog-based online journaling did not promote students' reflective thinking skills. Students who took part in solitary blogging showed a much higher level of reflection than students who engaged in peer feedback process.

On the other hand, the study conducted by Ertmer, Richardson, Belland, Camin, Connolly, Coulthard, et al. [16] indicated that although the students prefer instructor feedback, the quality of their postings in an online course was maintained through the use of peer feedback.

Meerah and Halim [19] examined the views of postgraduate students and academics with regard to feedback in learning. According to the results, the frequency of feedback received by students was low and of poor quality. Based on the students and academics' opinions on giving feedback, the researchers proposed a plan of action using feedback from peers. The experimental study showed that peer feedback had positive results and students' collaboration helped their learning process.

Cho and MacArthur [12] examined whether feedback from multiple peers improves the quality of students' writing more than feedback from the instructor. Further, it was tested how this feedback was associated with subsequent corrections and how these corrections affected the quality of students' writing. Undergraduate students had to complete written assignments as part of a compulsory course in research methods in psychology. After completing the task, the students received feedback from the instructor, a peer or multiple peers. The findings indicated that the students who received feedback from the teacher made simpler corrections than those who received peer feedback. Conversely, students who received feedback from multiple peers made more corrections. The overall conclusion of this research supports the positive influence of peer feedback in writing.

4 Synchronous vs. Asynchronous Communication

Synchronous communication refers to communication in real time and requires the simultaneous participation of respondents, such as the communication occurring in chat rooms. On the other hand, asynchronous communication concerns communication which does not require simultaneous participation of respondents, such as the communication occurring via electronic mail [20]. Thus, asynchronous communication is characterized by lack of immediate response, increased time required to complete a discussion and, potentially, a sense of isolation. At the same time, the time delay in completing a communication task allows more time to process the response. This encourages those who do not feel comfortable in engaging in a direct, synchronous conversation.

Synchronous communication is characterized by lack of time to process messages. Synchronous discussions are more difficult to take place than asynchronous ones since they are more demanding in terms of response time, knowledge and timing. However, synchronous discussions offer simultaneous communication among more users, and a greater sense of presence and spontaneity as it imitates interpersonal communication [20].

Several studies have investigated the use of synchronous and asynchronous communication in providing and receiving peer feedback. Kear [5] studied the different uses of asynchronous discussion in distance learning with distance learning students attending a UK course. Computer conferences were organized to provide peer support and assessment. The results showed that the practice of students offering support to their peers was effective. Students learned from one another through discussions, material exchange and by providing feedback on each other's work. The students also emphasized the importance of combining synchronous and asynchronous communication, especially at the beginning of the task, so that team members get to know each other.

Mabrito [6] studied a group of students working on essay writing and communicating via synchronous and asynchronous discussions. Groups of 3-5 students participated in the study and had to complete two tasks. For the first task, two of the groups were asked to use synchronous discussion via a chat tool. The other two groups were asked to use asynchronous discussion boards to collaborate. In the second task, the groups alternated their way of communication from synchronous to asynchronous and vice versa. Mabrito [6] concluded that when students communicated through synchronous discussions, they spent more time on communication. In addition, most of the synchronous discussions that took place focused on teamwork and dialogue. In the asynchronous communication situations, students communicated less but most of the discussion concerned the task. Thus, the researcher concluded that asynchronous communication can be the most effective way for written assignments which demand cooperation, even though asynchronous communication offered fewer opportunities for communication and intimacy among group members outside the boundaries of the task.

Last to be discussed, Park and Bonk [7] investigated the opportunities and challenges students had as they engaged in synchronous communication. The participants in this study were 22 distance learning students and 11 students who attended a face-to-face post-graduate course in educational technology. All students worked through the Breeze online system (which included written discussions, conferences and calls) to carry out the requested work; peer assessment was also facilitated via the synchronous tools offered by Breeze. Study findings indicated that the students appreciated the immediate feedback from peers, the interpersonal communication and the multiple perspectives on each subject. The regular communication with other team members enhanced the participants' social presence and sense of connection. Time constraints, lack of reflection, difficulty in communication relating to language (for foreign students), problems with the tool and internet connection were considered as challenges to synchronous communication.

5 Research Questions

The review of the literature indicated a research gap in this newly emerging field of peer feedback through synchronous and asynchronous communication tools. Specifically, there is little research investigating peer feedback, provided via synchronous and asynchronous communication, in face-to-face classroom environments. Thus, the purpose of the present study is to contribute to a better understanding of this issue by providing answers to the following questions:

1. What are undergraduate students' views regarding the use of peer feedback and the specific tools used for its technological mediation (synchronous vs. asynchronous collaboration)?
2. To what extent did the students perceive the peer feedback methods used as useful for communicating or learning?

6 Research Methodology

6.1 Participants

The participants were 35 undergraduate students of a university in Cyprus. Students used the online platform STOCHASMOS [21] for face-to-face activities during the 13 weeks of the course. Participation was voluntary but an extra bonus for participating in the study was provided. Eight groups, consisting of 4-5 people each, were formed based on their overall grade and gender, so that groups were of equal overall abilities.

Students engaged in group work, assigned by the instructor, and which related to the subject matter of each session. After completing each assignment, each group gave and received peer feedback, using one of two technology-mediated communication tools. Four groups used the STOCHASMOS in-context commenting to review each other's work (asynchronous communication), while the remaining four groups used the STOCHASMOS chat tool (synchronous communication).

6.2 The Online Teaching and Learning Platform STOCHASMOS

The STOCHASMOS online platform [21] is a web-based learning environment which supports students' reflective inquiry with complex, multi-modal data [22]. The students' learning environment consists of two main environments: the inquiry environment, in which the multi-modal data are organized, and the reflective WorkSpace, which is a separate area where students can organize their data in templates created by the teacher, and externalize their thinking, and engage in evidence-based reasoning [22]. STOCHASMOS provides tools that allow students to asynchronously share their WorkSpace pages with their peers and provide in-context commenting, and a chat tool for synchronous communication between paired groups.

6.3 Data Collection

Data collection lasted about three months, and included data from all students' individual pre- and post-questionnaires, the case study, and interviews.

Data Collection from all Participants. The 35 participants responded to pre-and post-course questionnaires, containing open and closed questions. The first questionnaire was completed at the beginning of the study and included general questions related to the students' views of feedback and its importance. The second questionnaire was administered at the end of the course and included questions concerning the students' experience with providing feedback and the tools used.

Data Collection from Small Group Interactions. To answer the question on the nature of reflection derived from peer feedback and its use to improve their work, eight students were randomly invited to participate in additional, small group activities and interviews. These students took part in a structured investigation and peer-review sessions and in individual, end-of-the-activity interviews.

The small group activities had students work in two groups to answer a problem presented in a new, online environment on STOCHASMOS. Two 3-hour sessions were held outside of regular class hours during the sixth week of classes. During the first session, the groups addressed the topic ‘New Technologies in Education’; during the second session students examined a topic titled ‘Distance Learning’. Each group was asked to provide their answers using the STOCHASMOS template pages, to share them with the other group and to provide feedback to each other. The groups were not located in the same room. During the first meeting, the participants were instructed to use the in-context commenting (Fig. 1). In the second meeting, participants were asked to use the chat tool to give and receive feedback (Fig. 2). All sessions were video-recorded.

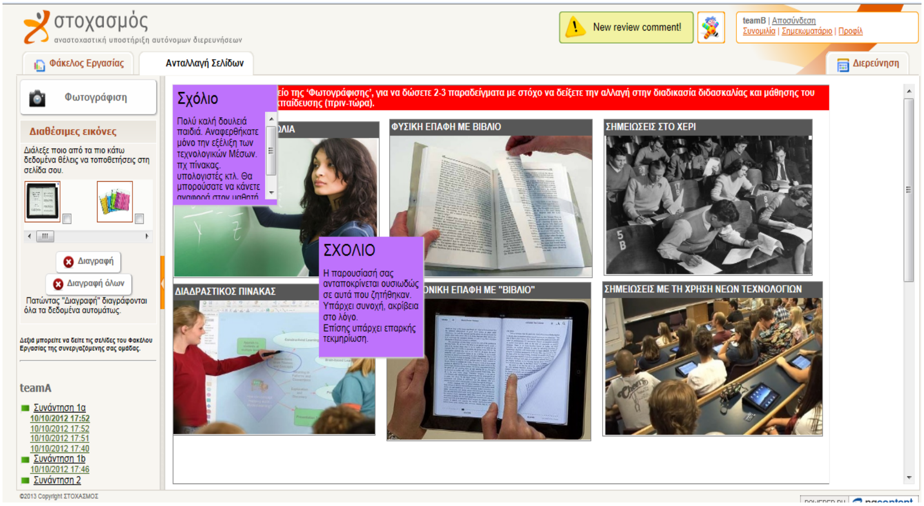


Fig. 1. In-context commentary (in Greek) in the STOCHASMOS reflective WorkSpace

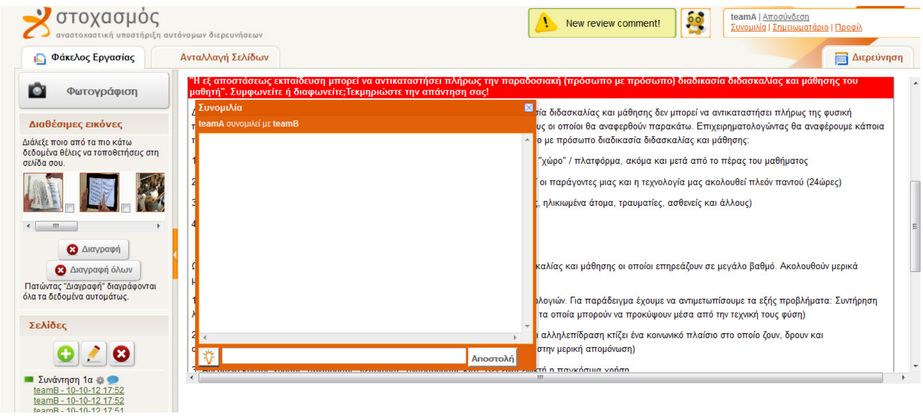


Fig. 2. The STOCHASMOS Chat tool (group work in Greek)

6.4 Data Analysis

After the completion of the data collection, coding and analysis of the results followed, based on the research questions and the data obtained. Tables of different categories reflecting each of the settings of data collection were created using grounded theory. The most representative examples of each category are presented in the sub-sections below. The first author and a graduate student engaged in inter-rater reliability testing, which was initially estimated at 91%, with all disagreements subsequently resolved between them.

7 Results

The results section is organized according to the two research questions; we first discuss students' reports on the importance of peer feedback and the specific tools they used, and then present results on the students' views of whether the tools supported their efforts to communicate and learn from each other.

7.1 Students' Views on the Importance of Peer Feedback

Figure 3 presents the participants' views on peer feedback before and after engaging in peer feedback. While at the outset of the study only 2,9% of the participants reported that peer feedback was very important to them, at the end of the study this percentage increased to 20,6%.

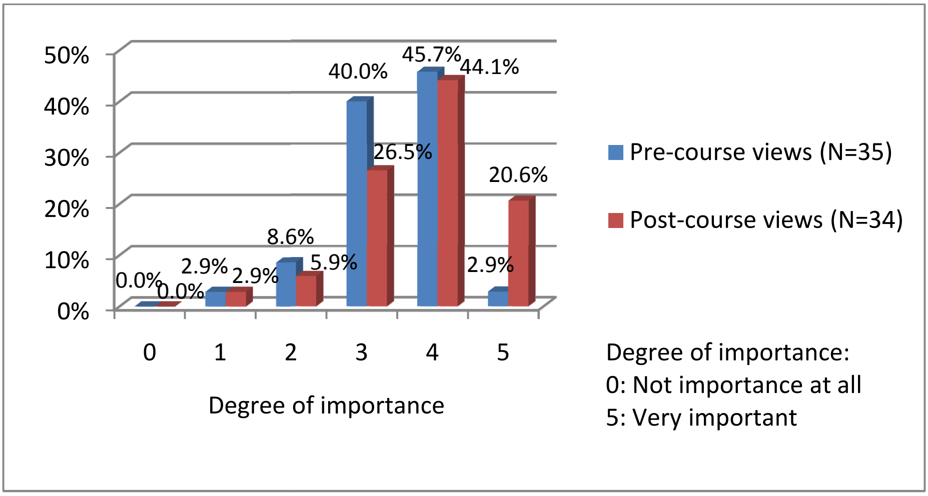


Fig. 3. Participants' views on the importance of peer feedback (pre- and post-course)

7.2 The Use of Peer Feedback

The analysis of the videotaped discussions in the small group settings indicated that the two case study groups noticed the peer feedback they received.

> *"Uh in cases we had to ... we considered that we had to change some things we changed them, but because many times we covered the topic it was not necessary [to make any changes] because we agreed with what the other team's members told us..."* (Markos, post, TeamA)

> (TeamA: Reads the comments received from TeamB)
> *"Very good wording of your ideas and we agree that although there are negative aspects too, in the end technology has offered a lot to learning."*
> TeamA: *"ok"*

Both groups seriously considered the comments of the other group and discussed several of them. This provides evidence that the groups took into account the feedback they received. As observed from the videos, as well as from the interviews, the students reported that there were instances when they modified their initial opinion based on comments received, if the team members agreed that what they initially wrote was indeed wrong.

> (TeamB reads TeamA's shared work...)
> (Discussion between the members of TeamB)
> Antonis: *"I disagree with the fear of the industrialization of education."*
> Manolis: *"Me too."*
> (TeamB writes comments for the other team)
> TeamA replies:
>
> > *"Respect… by the way, if you realized it, we agree that traditional methods are still present. We shall go on with our comment, regardless of the development of technological means. So we agree on this."*
>
> TeamB replies:
>
> > *"Yes, the only thing we don't agree is about industrialization of education. The overall opinion is the same but we disagree on some points that we may have perceived differently as we said on the industrialization of education."*
> >
> > *"Anyway we have already written that they will be changed but we don't agree that, for example, the computer will deliver the lesson. It's the person behind the computer, it is just more unnatural, this thing cannot undergo industrialization."*
>
> (Discussion between the members of TeamB)
> Anna: *"But we won't change our project, I don't agree."*
> Elena: *"Me neither."*

Thus, it was observed that the two groups did not only exchange feedback but also explained their comments. In the above excerpt, there was agreement between the members of the two groups, but, there was also reflection and discussion between them. On another case, although the issue was discussed, the two groups did not agree, so the members of the team (TeamB) did not modify their reply. In their interviews (Excerpt 1) students reported that they reflected on the feedback they received and went on to improve their work, if they felt that the comments received from peers were correct. In contrast, in Excerpt 2, students reported that if they did not agree with the comments received, they did not proceed with revising their work.

> Interview Excerpt 1: *"uh we read their comments and then we discussed them between us to see what we can answer, where we made mistakes to correct them. …sometimes we corrected some things."* (Vasso, post, TeamA)
>
> Interview Excerpt 2: *"First of all we read ours again and where they pointed a negative point, we read it again and we wondered if indeed they are right or not. Uh then based on our judgment first, if we decided that they were wrong we didn't change it."* (Anna, post, TeamB)

Studying students' views on the use of feedback, all of them reported that they took into consideration the feedback they received from their peers; they, then, either proceeded to changes to their work or not.

7.3 Comparing the Peer-Review Tools: In-Context Commenting vs. Chat

Tables 1 and 2 present students' perceptions of the positive and negative aspects of the tools used. The data were collected by open questions included in the post-course questionnaires completed by all participants which reported on their experience of the tool each of them used (chat or in-context commenting).

Table 1. Chat tool: Positive and negative aspects

	Chat
Positive aspects	• Immediate response • Friendly discussion • Comfortable communication • Fast problem solution • More appropriate for this age group • Quick process of feedback
Negative aspects	• Confusion between the groups' comments • Discussion focused on issues other than the task • Time consuming • Timing problems • Technical problems with STOCHASMOS • No comfortable expressing of thoughts • Tension during discussion

Table 2. In-context commenting: Positive and negative aspects

	In-context commenting
Positive aspects	• No need for both interlocutors to be connected at the same time • Completed comments • More careful processing of the comments given • No confusion on who said what • Always present on the platform • More focus on the subject
Negative aspects	• No response to the comments • Problem with understanding of the comments • Tension in interpersonal relationships • Technical problems with STOCHASMOS→ did not show others' comments • Delayed response or no response

7.3 Students' Views of the Contribution of In-Context Commenting and Chat to Their Learning

Regarding the extent to which the tools used helped the students in their learning, the following view of one of the case study students provides a representative answer.

> *"I think that in-context commenting is more useful in learning, because in chat you can write things that are not necessary, while in in-context commenting you write less and more precise things. The in-context commenting was better for learning and chat was better for communication."* (Antonis, post, TeamB)

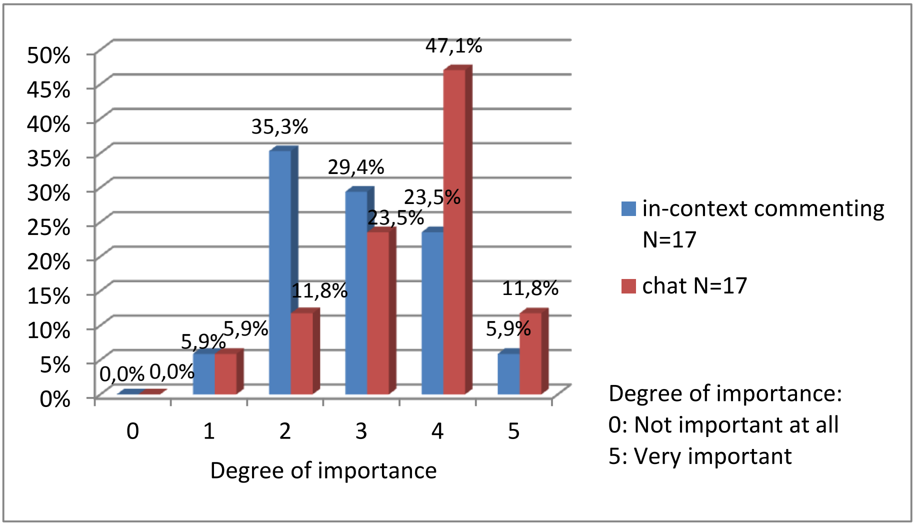


Fig. 4. Students' views of the importance of in-context commentary and chat in their communication

Respondents noted that the chat tool helped them to communicate better with each other, while the in-context commenting was best for their learning. This is also reflected in Figure 4 and Figure 5, which represent all 35 participants' views. Since half of the respondents to the post-class questionnaire (n=17) used the in-context commenting tool and half of them (n=17) used the chat tool to provide peer-feedback, the students' answers are grouped accordingly.

As Figure 4 shows, about 41% of the students who used the in-context commenting said that it did not help them communicate with each other; 29.4% had a neutral view, while 29.4% noted that it helped them a lot. On the other hand, a large percentage of respondents (58.9%) reported that the chat tool was very important for their communication, with only 17.7% of the students who used the chat tool arguing that it did not help them much in their communication.

As shown in Figure 5, 64.7% of the students who used in-context commenting said this tool helped them with their learning; a small percentage, 11.8%, reported that it did not help them, whereas 23.5% had a neutral view. On the other hand, 23.6% of the students who used the chat tool said that it helped them in their learning, 35.3% had a neutral view, while 41.2% noted that it did not help much.

To conclude, as indicated by students' responses, the majority of students who used the chat tool emphasized that the chat tool was useful for communication, whereas the majority of students who used the in-context commenting tool indicated that this communication facilitated their learning process.

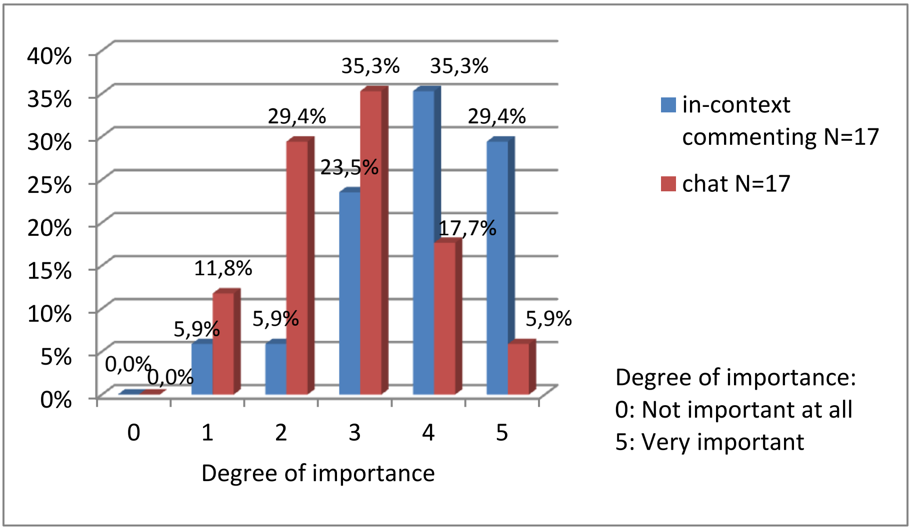


Fig. 5. Students' views of the importance of in-context commentary and chat in their learning

8 Discussion

Based on our findings, we can conclude that peer feedback can constitute an important process in the context of re-framing modern education. The results of our study showed that the role of students, when providing and receiving feedback, was not restricted to a passive one, but students had an active and productive participation in the learning process. Such involvement helped them to take initiative to evaluate the suggestions received from peers which implies the presence of critical thinking as Liu et al. [23] claimed. Consequently, it appears that both the provider and receiver of feedback can benefit as the peer-review process required them to read, compare, challenge ideas, suggest changes and evaluate their peers' work.

In the case of online feedback, we observed that students who used the chat tool enjoyed the use of synchronous communication, even though they realized that it can be a more time-consuming process that may not focus solely on the task. The technological tool used may play a key role in the whole process of feedback, since each one has its own affordances. A similar conclusion was drawn by Mabrito [6] who found that, in asynchronous communication, although students communicated less, most of the discussion focused on the task.

Another finding of the current study is the fact that all participants took into consideration the feedback they received. However, this might be indicative of socially desired responses. Such a finding is opposed to what Guardado and Shi [24] argued that the participants in their study took little account of their peers' comments and that some of them refused to take responsibility to express and explain comments. This might have happened because the researchers examined students' experiences about online peer feedback in a different context than the one in the Guardado and Shi's study (Learning English as a Second Language). Therefore, the subject on which to give or receive feedback may be a determining factor in this process. Future research may examine peer feedback experiences of students in different disciplinary areas. Larger sample sizes and purposive sample of population of maximum diversity (in grades, skills, etc.) can help test the findings of this study.

To sum up, peer feedback via synchronous and asynchronous communication can be considered an important activity in the context of face-to-face higher education since, as this study demonstrated, students recognize benefits from this process. As a result, one can argue that the use of technologically mediated communication, and especially of asynchronous communication, in face-to-face courses, can bring positive results to the teaching and learning process.

Acknowledgements. The data presented in this paper are part of the master's thesis work conducted by the first author, and supervised by the second author, during the first author's studies for her MA degree in the program "New Technologies for Communication and Learning" in the Department of Communication and Internet Studies at the Cyprus University of Technology. The authors would like to thank all the students who participated in this study, as well as Nicolas Finopoulos and Raphael Sofokleous for their help with data collection.

References

1. Hattie, J., Timperley, H.: The power of feedback. Review of Educational Research 77(1), 81–112 (2007)
2. Topping, K.: Peer assessment between students in colleges and universities. Review of Educational Research 68(3), 249–276 (1998)
3. Mory, E.H.: Feedback research revisited. In: Handbook of Research on Educational Communications and Technology, pp. 745–783 (2004)
4. Falchikov, N.: Learning Together: Peer Tutoring in Higher Education. Routledge Falmer, London (2001)
5. Kear, K.: Peer learning using asynchronous discussion systems in distance education. Open Learning: The Journal of Open, Distance and e-Learning 19(2), 151–164 (2004)
6. Mabrito, M.: A study of synchronous versus asynchronous collaboration in an online business writing class. The American Journal of Distance Education 20(2), 93–107 (2006)
7. Park, Y., Bonk, C.: Synchronous learning experiences: Distance and residential learners' perspectives in a blended graduate course. Journal of Interactive Online Learning 6(3), 245–264 (2007)
8. Mory, E.H.: The use of informational feedback in instruction: Implications for future research. Educational Technology Research and Development 40(3), 5–20 (1992)
9. Matsangouras, E.: Strategies of Teaching. Gutenberg, Athens (1998)
10. Gibbs, G., Simpson, C.: Conditions under which assessment supports students' learning. Learning and Teaching in Higher Education 1(1), 3–31 (2004)
11. Kalaitzi, E., Derri, V., Vasiliadou, O., Kioumourtzoglou, E.: Academic learning time in elementary Physical Education class. Inquiries in Sport & Physical Education 5(2), 187–198 (2007)
12. Cho, K., MacArthur, C.: Student revision with peer and expert reviewing. Learning and Instruction 20(4), 328–338 (2010)
13. Born-Lechleitner, I., Kepler, J.: Peer feedback in blended learning. Vienna English Working Papers 18(3), 185–187 (2009)
14. Corgan, R., Hammer, V., Margolies, M., Crossley, C.: Methods-making your online course successful. Business Education Forum 58(3), 51–54 (2004)
15. Falchikov, N.: Improving learning through critical peer feedback and reflection. Different approaches: Theory and practice in Higher Education. In: Proceedings of HERDSA Conference (1996)
16. Ertmer, P., Richardson, J., Belland, B., Camin, D., Connolly, P., Coulthard, G., et al.: Using peer feedback to enhance the quality of student online postings: An exploratory study. Journal of Computer Mediated Communication 12(2), 412–433 (2007)
17. Liu, N., Carless, D.: Peer feedback: The learning element of peer assessment. Teaching in Higher Education 11(3), 279–290 (2006)
18. Xie, Y., Ke, F., Sharma, P.: The effect of peer feedback for blogging on college students' reflective learning processes. Internet and Higher Education 11, 18–25 (2008)
19. Meerah, T., Halim, L.: Improve feedback on teaching and learning at the university through peer group. Procedia - Social and Behavioral Sciences 18, 633–637 (2011)
20. Johnson, G.: Synchronous and asynchronous text-based CMC in educational contexts: A review of recent research. TechTrends 50(4), 46–53 (2006)

21. Kyza, E.A., Constantinou, C.P.: STOCHASMOS: A web-based platform for reflective, inquiry-based teaching and learning. In: Cyprus: Learning in Science Group (2007), http://www.stochasmos.org
22. STOCHASMOS: Web-based reflective inquiry environments (2012), http://www.stochasmos.org/nqcontent.cfm?a_id=792
23. Liu, E.Z.F., Lin, S.S.J., Chiu, C.H., Yuan, S.M.: Web-based peer review: The learner as both adapter and reviewer. IEEE Transactions on Education 44(3), 246–251 (2001)
24. Guardado, M., Shi, L.: ESL students' experiences of online peer feedback. Computers and Composition 24(4), 443–461 (2007)

Virtual Labs in Engineering Education: Modeling Perceived Critical Mass of Potential Adopter Teachers

Raghu Raman, Krishanshree Achuthan, and Prema Nedungadi

School of Engineering, Amrita Vishwa Vidyapeetham,
Amritapuri, Kollam, Kerala, India - 690525
{raghu,krishna,prema}@amrita.edu

Abstract. Virtual labs for science experiments are a multimedia technology innovation. A possible growth pattern of the perceived critical mass for virtual labs adoption is modeled using (N=240) potential-adopter teachers based on Roger's theory of diffusion and of perceived attributes. Results indicate that perceived critical mass influences behavior intention to adopt a technology innovation like Virtual Labs and is affected by innovation characteristics like relative advantage, ease of use and compatibility. The work presented here models the potential-adopter teacher's perceptions and identifies the relative importance of specific factors that influence critical mass attainment for an innovation such as Virtual Labs.

Keywords: virtual labs, innovation diffusion, critical mass, simulation, lab experiments.

1 Introduction

Virtual learning environments like Virtual Labs have played a catalytic role in improving conceptual learning across the higher education landscape. The trends and challenges elicited in Mueller's work [1] highlighted an important factor affecting its implementation i.e. the perception of such an environment by the users and therefore its rigorous usage.

The dominant focus of the content coverage of virtual learning environments has been on theory concepts. While this is important, practical knowledge on usage and applications of theory is equally critical in science and engineering curriculum [2], [8]. Cognizance of fundamentals, conceptual understanding of its applications in the field of study, sensory attentiveness, psychomotor skills development etc. forms the gamut of objectives for laboratory exercises. The limited experimental infrastructure and restricted time in most institutions make it challenging to ensure all students attain these learning outcomes in a satisfactory fashion.

Recent years have seen creation of a new dimension to virtual learning with the introduction of virtual labs. Virtual labs have been augmenting the hands-on physical laboratories in several ways by allowing students to learn most elements of the experiments prior to physically seeing or experimenting with the instrumentation [9], [14]. Physical laboratories group students to perform activities while virtual labs

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 288–300, 2013.

provide individualized training. Virtual labs are in fact preferred in many educational scenarios for scalability and tacitly addressing fundamental issues such as the inability to visualize complex phenomena, compromise in topical coverage of the content due to variations in teaching styles etc. In spite of its advantages of the academic community hasn't integrated virtual labs in a global way as is the case with physical labs. Unless there is massive adoption of virtual laboratories, its large-scale impact will remain a distant reality.

Many technological innovations have required espousal by a threshold 'critical mass' for their main-stream usage and sustenance [10]. Some examples include: E-platforms used for online businesses [12] showed success when factors such as user's decision and user network surpassed the critical mass. Grajek [11] developed empirical models in global cellular telephony that predicted the critical mass based on heterogeneous factors and showed how the timing of technology introduction is critical. Grewal [13] emphasized that the depth of usage of an innovation apart from the critical mass should also be considered in ascertaining the level of adoption.

The phenomenon of critical mass is important in the context of virtual labs due to the interdependent and collaborative aspect of technology adoption in teaching processes. Rogers [10] defines critical mass as 'the point after which further diffusion becomes self sustaining.' Loch and Huberman [19] concluded that critical mass can be attained if users have a high rate of trialability. The threshold level for critical mass varies anywhere between 10% [20] and 25% [21]. Our work analyzes the potential adopter teacher's intention to adopt virtual labs using their perception of critical mass. Thus it becomes important to consider 'how many' should adopt virtual labs for it to become a conventional technology.

Starting with a research model we develop various hypotheses and describe the research methodology followed to empirically validate the hypothesis with conclusions and recommendations.

2 Literature Survey

2.1 Virtual Labs for Engineering Education

Review of the literature on physical and virtual labs by Ma et. al [5] showed lack of foundation to evaluate effectiveness of laboratories and the need to consider learning outcomes and student preferences more closely in an isolated fashion. Corter et al [4] compared the effectiveness of physical laboratories with remotely controlled and simulation based virtual laboratories and showed that all three were effective in terms of conceptual understanding they imparted to the students. Today virtual laboratories that are simulation-based or remotely triggerable have been developed for physical & chemical sciences, engineering, biotechnology and medical sciences. The VALUE labs (Virtual Amrita Laboratories Universalizing Education) [14] has built over 400 experiments that are interactive animations, mathematical simulations of physical phenomena and interfaced a variety of equipments as part of its virtual laboratory development program. This was developed as part of larger consortia of 12 institutes whose mission was to develop over 1500 experiments in nine disciplines of engineering and sciences (http://vlab.co.in). The concept of 'discipline wise national coordinator' (DNC) was introduced to assign the responsibility of coordinating the virtual lab development of a particular discipline to a specific consortia partner. The DNC for a discipline made sure virtual laboratory experiments aligned to the laboratory curriculum of that discipline.

2.2 Critical Mass of Potential Adopters

Rogers [10] suggested that technology adoption depended on the mental attitude of the adopters and classified them as 'innovators', 'early or late majorities' and 'laggards' based on their extent of adoption. The concept of 'critical mass' and endogenous growth of adoption that Rogers described has been successfully demonstrated in many applications [13,16]. The social influences, networking with positive feedback looping in the communication exchanges accelerate attainment of critical mass required to justify adoption [15]. The diffusion curve starts to even out when approximately half of the individuals in a social system have adopted the innovation, defined as a 'critical mass'. This is because every new adopter finds it increasingly difficult to convey the idea to a peer that has yet to adopt, as individuals that are unaware of the innovation become increasingly scarce. Critical mass is achieved when enough individuals have adopted the innovation so that it becomes self-sustaining. Grewal's findings [13] outlined for critical mass to be achieved in the use of a new online learning platform, user support and training are necessary in facilitating adoption of the new system.

Wang et. al [7] model the factors affecting instructor's attitude and intention to use web-based learning systems. Their work showed that increased perceived usefulness also increased the intention for use. Reflecting on the challenges with adoption of virtual laboratories, issues working against its prevalence may be technological and/or pedagogical in nature. Factors most considered in such a debate are lack of infrastructure, internet bandwidth, proficiency in use of technology, attitudinal variations of teachers and poor human-computer interactivity [3]. Instructors play a very critical role in determining the success or failure of virtual learning systems. Yoon et. al [6] discuss the lack of adoption of virtual reality worlds by organizations from a technology-organization-environment framework and conclude that this is partially due to lack of other organizations adopting it. In other words, it is important for competing organizations to adopt a new technology for institutional isomorphism leading to the rapid diffusion of innovation.

3 Case study: Modeling Perceived Critical Mass of Potential-Adopter Teachers

3.1 Research Model and Hypothesis

In the diffusion model of Virtual labs we chose teacher perception of critical mass as the dependent variable and two groups of characteristics - innovation characteristics and environment characteristics as independent variables (Fig. 1).

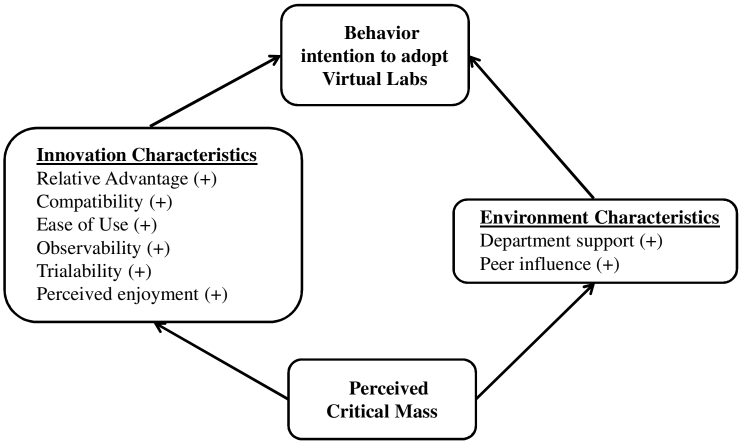


Fig. 1. Research model for perceived critical mass on Virtual Labs adoption

3.2 Innovation Characteristics

Relative Advantage. Rogers [10] defines Relative Advantage as — 'the degree to which an innovation is perceived as being better than the idea that it supersedes'. Here the focus is on the potential adopter and his perception of advantages of the innovation and not so much on the advantages proposed by the producer. We hypothesize that

H1:Perceived critical mass will have a positive effect on teacher's beliefs about the Relative Advantage of Virtual Labs

Compatibility. Rogers [10] defines compatibility as — 'the degree to which an innovation is perceived as consistent with existing values, past experiences, and needs of potential adopters'. We hypothesize that

H2:Perceived critical mass will have a positive effect on teacher's beliefs about the compatibility of Virtual Labs

Complexity. Any innovation quickly gains a reputation as to its ease or difficulty of use [10]. An important question is to what extent Virtual Labs are perceived by users as simple to use. The idea of complexity, as described by Rogers' [10], was formulated from an "Ease of use" perspective in this study whereas the notion of adoption was substituted with the notion of attitude towards use. We hypothesize that

H3:Perceived critical mass will have a positive effect on teacher's beliefs about the ease of use of Virtual Labs

Trialability. This is 'the degree to which an innovation may be experimented with on a limited basis' [10]. Innovations that potential adopter can play with on a trial basis are more easily adopted because an innovation that can be tried presents less risk to the potential adopter. If there exist a way and opportunity to try an innovation and acquire personal experiences, it will notably decrease user apprehensions towards that innovation. We hypothesize that

H4:Perceived critical mass will have a positive effect on teacher's beliefs about Trialability of Virtual Labs

Observability. Another aspect of Rogers theory [10] of perceived attributes is related to —the degree to which the results of an innovation are visible to others. Rogers writes [10] — 'the observability of an innovation, as perceived by members of a social system, is positively related to its rate of adoption'. If potential adopters can see the benefits of an innovation, they will easily adopt it. We hypothesize that

H5:Perceived critical mass will have a positive effect on teacher's beliefs about the observability of Virtual Labs

Perceived Enjoyment. According to Rogers [10] Perceived enjoyment is the 'degree to which using an innovation is perceived to be enjoyable in its own right and is considered to be an intrinsic source of motivation'. The enjoyment that is perceived to be derived by using Virtual Labs is, therefore, expected to affect the attitude and the intention of teachers to adopt them. We hypothesize that

H6:Perceived critical mass will have a positive effect on teacher's beliefs about the perceived enjoyment of Virtual Labs

Department Support. More often teachers are motivated to consider technology decisions that are sanctioned by the department management since those will have adequate support resources for training, infrastructure etc. We hypothesize that

H7: Perceived critical mass will have a positive effect on teacher's beliefs about the departmental support for Virtual Labs

Peer Influence. Exchange of ideas occurs much more frequently between potential-adopters who are peers. In this regard the peer's perception of the innovation will strongly influence teacher's thinking. We hypothesize that

H8: Perceived critical mass will have a positive effect on teacher's beliefs about their peer's influence about Virtual Labs

3.3 Research Methodology

Sample and Method of Dissemination for Virtual Labs Teacher Workshops. A database of institutes, department heads was created with their contact numbers and emails. The target institutes had never used computer mediated course materials for any of their laboratory courses. A brochure eliciting an introduction to virtual labs importance of virtual laboratories, the content coverage with location, date and time was created. Following this, a four-prong approach to advertising was chosen as shown in Figure 2. Brochures were sent to the head of the institutes and the departmental head by postal mail. Calls were made to the department heads and a verbal affirmation or decline of attendance was received. In case of institutes that were within 150 km area, visits were made to personally invite faculty. 243 teachers (male = 97, female = 146) participated through four different workshops.

Fig. 2. Publicity approach to the workshops

Structure of Workshops. The workshop was designed to cater for all levels of IT literacy from the novice to advance users. The whole program was engineered to hold participants interest, to evoke curiosity, and to excite them about the potential and possibilities of virtual laboratories. The process was a step-by–step guide starting with the basics defining the concepts of virtual labs leading to presentation of topical areas, followed by demonstration of experiments with interactive discussions and culminating with hands-on training.

The workshop had six main parts:

Registration. At the beginning of the workshop a registration process was required and faculty background information was collected. It was found that most teachers registered online. Workshop attendee profile is shown in Fig. 3.

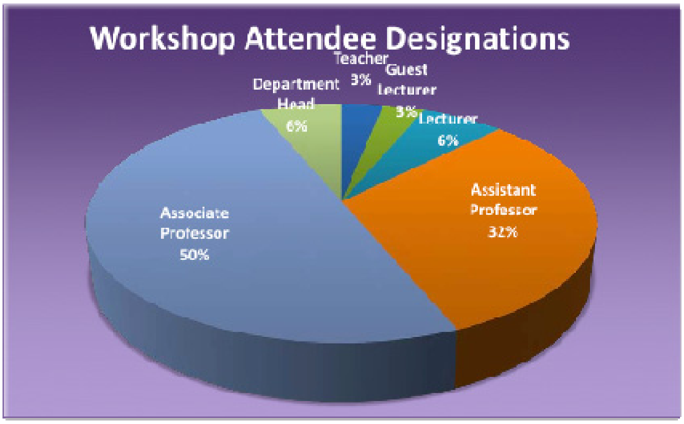


Fig. 3. Workshop attendee profile

Introductory Session. Most teachers had never heard of the virtual labs program prior to attending the workshop. The vision, objectives, methodologies and current state of development was first presented to them. This was followed by the elaboration of the advantages of using Virtual Labs based on the feedback received. This introductory session was common to teachers from different engineering streams.

Demonstration of Virtual Lab Experiments. Teachers were divided into groups pertaining to their field of expertise. Subject matter experts that developed Virtual Labs content guided the demonstration section. A list of topics covered under the theme of physics, chemistry or biotechnology was presented. Specific experiments were taken for demonstration due to shortage of time. The selected experiments belonged to various classes of virtual experiments i.e. interactive animations (for protocol intensive experiments), simulation based (mathematically modeling physical phenomena) and remote triggered experiments (equipments interfaced via the web for remote experimentation). This session was made interactive so that teachers are involved in questioning the various features and their utility.

A typical Virtual Lab experiment with theory, procedure, animator, simulator, assessment etc. is shown in Fig. 4.

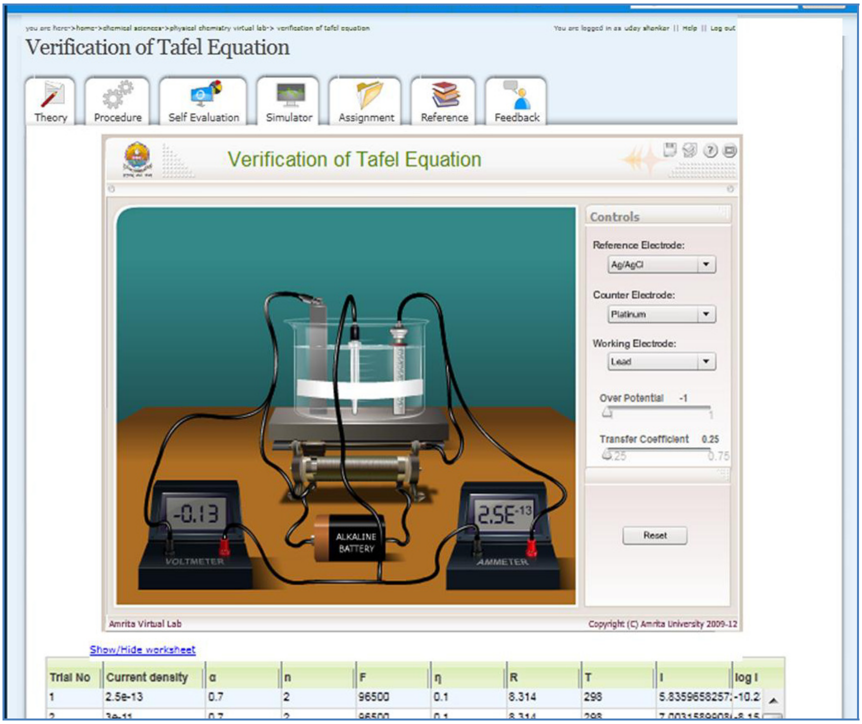


Fig. 4. Sample Virtual Lab experiment

Hands on Training. For the hands-on training, teachers were given a dedicated computer. Instructions were given on how to log into the system. Once they logged in, they were allowed to explore any experiment. Guidance was given at an individual level if and when questions were raised. Most teachers completed several experiments to check the accuracy and realism behind the experiment.

Feedback. After completion of the experiments, a detailed questionnaire was filled out by every one.

Nodal Center Forum. The concept of nodal centers was introduced the attendees. Nodal center program is an exciting new venture which allowed faculty and their institutes to actively partner in the development and implementation of virtual laboratories. A commitment to usage of virtual labs by provisioning computer labs and exposing students and/or other institutes in their periphery were the only requirements to becoming a nodal center. Other advantages of being nodal center included faculty following the progress of Virtual Labs, providing a platform for everyone to contribute towards the future development of laboratories and experiments and networking with several other institutions.

4 Results, Analysis and Discussion

SPSS and R were used to analyze the data. Some innovation attributes emerged as dominant and more relevant to the behavior intention under study. In this section we do a systematic testing of the various hypotheses starting with reliability, discriminant and convergent validity analysis.

In our study reliability of the eight attributes had values ranging from 0.88 to 0.93. (Table 1). According to Nunnally (1978) for internal consistency, reliability Cronbach Alpha values of 0.70 and above is acceptable.

Table 1. Reliability Analysis

Variables	Attributes	# of items	Cronbach's Alpha
Innovation characteristics	Relative Advantage	8	0.92
	Compatibility	3	0.88
	Ease of Use	3	0.91
	Trialability	2	0.93
	Observability	2	0.89
	Perceived enjoyment	2	0.93
Environment characteristics	Department support	8	0.91
	Peer influence	4	0.89

In the Table 2, diagonal elements (bold) represent the Average Variance Extracted (AVE) between the attributes. The numbers below the diagonal elements are the shared variances) among attributes. For discriminant validity, diagonal elements should be larger than off-diagonal elements.

Convergent validity was analyzed by examining the average variance extracted (AVE) for each attribute. The AVE for an attribute reflects the ratio of the attribute's variance to the total variances among the items of the attribute. It is important to note that the AVE's are well above the recommended 0.50 level. (Table 3)

Multiple regression analysis was performed using all the 8 independent variables on the dependent variable perceived critical mass. Results are summarized in the Table 4.

There is strong support for Hypothesis H1 (Relative Advantage), H2 (Compatibility), H3 (Ease of Use), H6 (Perceived Enjoyment), and H7 (Department Support). The regression model was statistically significant ($p < .0001$) and accounting for 79% of the variation in perceived critical mass towards adoption of Virtual Labs ($\mathbf{R^2} = .79$).

Table 2. Discrimant Validity analysis

	RA	CO	EOS	TR	OB	PE	DS	PI
Relative Advantage (RA)	**0.94**							
Compatibility (CO)	0.56	**0.89**						
Ease of Use (EOS)	0.45	0.54	**0.88**					
Trialability (TR)	0.42	0.49	0.33	**0.90**				
Observability (OB)	0.45	0.43	0.39	0.44	**0.91**			
Perceived Enjoyment (PE)	0.55	0.51	0.45	0.49	0.36	**0.88**		
Department support (DS)	0.46	0.39	0.41	0.39	0.41	0.35	**0.92**	
Peer Influence (PI)	0.33	0.29	0.34	0.37	0.39	0.42	0.38	**0.88**

Table 3. Convergent Validity analysis

Attributes	**CR**	**AVE**
Relative Advantage (RA)	0.85	0.81
Compatibility (CO)	0.84	0.78
Ease of Use (EOS)	0.86	0.67
Trialability (TR)	0.81	0.69
Observability (OB)	0.85	0.71
Perceived Enjoyment (PE)	0.88	0.79
Department support (DS)	0.89	0.81
Peer Influence (PI)	0.85	0.78

Table 4. Summary of Hypothesis results

Attributes		**Beta**	**t-values**	**Result**
Relative Advantage (+)	H1	0.38	2.89	Accepted
Compatibility (+)	H2	0.31	2.62	Accepted
Ease of Use (+)	H3	0.34	2.64	Accepted
Trialability (+)	H4	-0.17	-1.22	Rejected

Table 4. (*Continued*)

Attributes		**Beta**	**t-values**	**Result**
Observability (+)	H5	-0.30	-1.31	Rejected
Perceived Enjoyment (+)	H6	0.32	2.82	Accepted
Department support (+)	H7	0.36	2.14	Accepted
Peer influence (+)	H8	-0.16	-1.23	Rejected
R^2 **Adjusted R^2**	0.79 0.75			

The largest beta coefficient is 0.38 which is the Relative Advantage of Virtual Labs which make this variable the strongest contributor in explaining the dependent variable (perceived critical mass), when the variance explained by all other predictor variables in the model is controlled for. It is equally interesting to see that the beta coefficient for Department Support is also near high confirming the importance of environment support for the adoption of Virtual Labs by the teachers. The predictor variables like Observability and Peer Influence were insignificant in their contribution.

5 Conclusions

Virtual Labs is still an emerging learning paradigm in laboratory engineering education. Our current study focused on potential-adopter teacher's perception of critical mass for Virtual Labs and ranked attributes like Relative Advantage, Ease of Use, Perceived Enjoyment and Department Support as dominant factors. Surprisingly enough Peer Influence did not come across as a significant factor. Critical mass refers to 'the point at which enough individuals have adopted an innovation so that the innovation's further rate of adoption becomes self-sustaining' [10]. Since it is difficult to pinpoint the exact timing of critical mass attainment, a potential-adopter's perception of whether it has reached or how much more time the innovation will take to attain critical mass is very helpful. This work emphasizesthat the critical mass of potential-adopter teachers is of paramount importance for sustaining a technology innovation like Virtual Labs. From a practical point of view, the findings of this study can be used by educational policy makers and higher educational institutions to establish or improve their virtual learning environments like Virtual Labs.

The outcome of this study is limited by time since the adoption decision process, according to Rogers [10] is not an instantaneous decision but rather an evolutionary one. Future studies could focus on the worldwide adoption of Virtual Labs to study potential influence of any demographic, cultural differences.

Acknowledgements. This project derives direction and ideas from the Chancellor of Amrita University, Sri Mata Amritanandamayi Devi. This work is in part funded by Amrita University and National Mission on Education through ICT (NMEICT), Govt.

of. India. The authors would like to acknowledge the contributions of faculty and staff at Amrita University whose feedback and guidance was invaluable.

References

1. Mueller, D., Strohmeier, S.: Design characteristics of virtual learning environments: state of research. Computers & Education 57, 2505–2516 (2011)
2. Zelby, L.W.: Engineering Approach to Engineering Education. IEEE Transactions on Education E-11(2), 89–93 (1968)
3. Kara, A., Ozbek, M.E., Cagiltay, N.E., Aydin, E.: Maintenance, sustainability and extendibility in virtual and remote laboratories. Procedia - Social and Behavioral Sciences 28, 722–728 (2011)
4. Corter, J., Nickerson, J.V., Esche, S.K., Chassapis, C.: Constructing Reality: A Study of Remote, Hands-on and Simulated Laboratories. ACM Transactions on Computer-Human Interaction 14(2), Article 7 (2007)
5. Ma, J., Nickerson, J.V.: Hands-On, Simulated, and Remote Laboratories: A Comparative Literature Review. ACM Computing Surveys 38(3), Article 7 (2006)
6. Yoon, T.E., George, J.F.: Why aren't organizations adopting virtual worlds. Computers in Human Behavior 29, 772–790 (2013)
7. Wang, W.T., Wang, C.C.: Am empirical study of instructor adoption of web-based learning systems. Computers & Education 53, 761–774 (2009)
8. Gustavsson, I., Garcia-Zubia, J., Nafalski, A., Pettersson, M.I.: On Objectives of Instructional Laboratories, Individual Assessment and Use of Collaborative Remote Laboratories. IEEE Transactions on Learning Technologies 2(4), 263–274 (2009)
9. Cooper, M., Ferreira, J.M.M.: Remote Laboratories Extending Access to Science and Engineering Curricular. IEEE Transactions on Learning Technologies 2(4), 342–353 (2009)
10. Rogers, E.: Diffusion of Innovations, 5th edn. Free Press, New York (2003)
11. Grajek, M., Kretschmer, T.: Identifying critical mass in the global cellular tele-phony market. International Journal of Industrial Organization 30, 496–507 (2012)
12. Huang, O., Duan, W.: Evolution Model and Critical Mass of E-business Plat-form Based on Complex Networks. In: Fifth Intl. Conference on Business Intelligence and Financial Engineering (2012)
13. Grewal, S.K., Harris, L.J.: Beyond Critical Mass: A Case Study Investigating the Use of WebCT for Course Delivery by Faculty in a Campus Based UK University. MERLOT Journal of Online Learning and Teaching 4(3) (2008)
14. Achuthan, K., Sreelatha, K.S., Surendran, S., Diwakar, S., Nedungadi, P., Humphreys, S., Sreekala, S., Pillai, Z., Raman, R., Deepthi, A., Gangadharan, R., Appukuttan, S., Ranganatha, J., Sambhudevan, S., Mahesh, S.: IEEE Global Humanitarian Technology Conference Proceedings, pp. 117-121 (2011)
15. Peng, G.: Critical Mass, Diffusion Channels, and Digital Divide. Journal of Computer Information Systems, 63–71 (2010)
16. Corigliano, M.A., Baggio, R.: Mobile Technologies Diffusion in Tourism: Mod-elling a Critical Mass of Adopters in Italy. In: 11th Intl. Conf. on Information Technology for Travel and Tourism (2004)
17. Virtual reality in the real world: History, applicationsand projections 25(9), 963–973 (November)

18. Jacobsen, D.M.: Examining Technology Adoption Patterns by Faculty in Higher Education. In: Proceedings of Ace 2000, Learning Technologies, Teaching and the Future of Schools, Melbourne, Australia (2000)
19. Loch, C., Huberman, B.: A punctuated-equilibrium model of technology diffu-sion. Management Science 45(2), 160–177 (1999)
20. Mahler, A., Rogers, E.: The diffusion of interactive communication innovations and the critical mass: The adoption of telecommunications services by German Banks. Telecommunications Policy 23, 719–740 (1999)
21. Cool, K., Dierickx, I., Szulanski, G.: Diffusion of innovations within organizations: electronic switching in the bell system, 1971–1982. Organization Science 8(5), 543–559 (1997)
22. Markus, L.: Toward a 'critical mass' theory of interactive media: universal access, interdependence and diffusion. Communication Research 14, 491–511 (1987)

Appendix A – Virtual Lab (VL) Survey Questionnaire

Survey options:

1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree

Relative Advantage

1. Using VL enables me to accomplish my teaching tasks more quickly
2. Using VL improves the quality of my teaching
3. Using VL enhances my teaching effectiveness.
4. Using VL gives me greater control over my teaching
5. VL supports student learning in different ways
6. I do not see too much need to utilize VL in the instructional process.
7. I do not see too much advantage of VL over thetraditional approach in the instructional process.
8. Classroom management is more difficult when using VL in teaching

Compatibility

1. Using VL will fit well with the way I teach and assess students
2. Using VL will fit into my style of teaching
3. I feel already over-burdened without adding VL into my instructional process

Perceived ease of use

1. Using VL to support my teaching is clear and understandable.
2. When using VL to support my teaching, it is easy to get tools to support VL
3. Overall, I believe that it is easy to use VL to support my teaching

Trialability

1. It is easy to recover from mistakes when using VL
2. Before deciding whether to use VL, I would want to be able to properly try them out.

Observability

1. I learned new ways of using VL from my colleague
2. I changed my way of using VL based on what I have learned from my colleague

Perceived enjoyment

1. The actual process of using the VL is pleasant
2. I enjoy learning new ways of teaching, assessing like VL for my students

Department support

1. The department understands the importance of using VL and encourages me to use them.
2. The department is committed to supporting my efforts in using VL for teaching
3. The department will recognize my efforts in using VL for teaching.
4. The use of VL for teaching is important to the department
5. I have pressures from my department to use VL
6. My department does not provide convenient time for getting trained on VL
7. I am happy with the procedures in my department to adopt innovations like VL
8. The technical support for innovations like VL in my department is satisfactory.

Peer influence

1. My teacher colleagues think that using VL is valuable for teaching
2. My teacher colleagues' opinions are important to me
3. If most of my colleagues have started to use VL to support their teaching, this fact would encourage me to do the same.
4. I will learn how to use VL after seeing my teacher colleagues use it.

Qualitative and Quantitative Evaluation of an Adaptive Course in GALE

Vinicius Ramos[1], Paul De Bra[2], and Geraldo Xexéo[1]

[1] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil
{vfcr,xexeo}@cos.ufrj.br
[2] Department of Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands
debra@win.tue.nl

Abstract. This paper presents and discusses in detail the results obtained by an evaluation of an adaptive course served by the GALE system developed in the EU FP7 project GRAPPLE. The main goal of that evaluation is to understand the influence of adaptation on students' learning in an adaptive hypertext course. We analyzed the students' logs, the performed tests and assignment grade as part of the quantitative method. The qualitative method consists of the analysis of a questionnaire, where students answered questions about their first experience in an adaptive course. Through the questionnaire, usage log, tests performed and assignment grades, we show that the students follow, in most of the cases, the adaptive structure of the course. We also present results about how the students felt about the adaptive course. At the end we discuss future work, and in particular suggest changes to the setup and adaptation in the course based on the observed student behavior as well as the student feedback.

1 Introduction

Adaptive Hypermedia (AH) systems, first summarized by Brusilovsky in 1996 [1] and in 2001 [2], and updated by Knutov et al. in 2009 [10], are systems that build an individual model for each user and apply it for adapting the system to that user. The individual model is based on his preferences, goals and knowledge. AH systems are used on the Web in the educational field to improve the students' experience in studying an on-line course. The course content can consist of information, assignments, activities and tests. The adaptation intends to, in most of the cases, offer the students freedom in the possible ways to study a course, without running into trouble that could be caused by studying parts that can only be understood after studying some other (prerequisite) parts. The GALE system was recently introduced as an evolution of the AHA! system [7,8]. It was created to support the main adaptive techniques that exists nowadays [12] and to allow extending the system to suit particular adaptation needs. The two main technique used in the adaptive courses currently served through GALE are: the adaptation engine can present links in different colors depending on whether

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 301–313, 2013.

it considers the student ready to study the link destination (page), and it can conditionally include fragments and objects in the page, for instance to show a prerequisite explanation. Link colors or colored icons suggest which links are relevant to each student; this approach dates back to early adaptive systems such as ELM-ART [4].

In the educational context, the past few years the number of AH systems has grown. The adaptive link annotation technique, has been shown to be an important tool to help students to acquire knowledge faster and easier, and to reduce navigation overhead, while at the same time moving away from the traditional sequential way of studying (typical for textbooks) [2,5,6]. Good surveys about the state of art of Adaptive Educational Hypermedia (AEH) systems can be found in [3,11]. A more specific work about evaluations of AEH systems is presented in [9].

In this paper we present in detail the results obtained by an evaluation of (the adaptation and navigation in) an adaptive course offered through GALE. We briefly describe the adaptive course in Sect. 2. The evaluation is based on quantitative and qualitative methods. We use data obtained from the access log, test logs, questionnaire, and exam grades from bachelor students at the Eindhoven University of Technology. In Sect. 3 we describe the results obtained by our evaluation. To understand the way users navigate through an adaptive course, our goal is to identify which pages and links influence the choices of the students and to contrast this with the test logs, questionnaire answers and exam grades as well as the adaptation rules employed by the adaptation engine. Another goal is to verify how the adaptation rules, created by the author of the course, influence the navigation by the students. The discussion about the results and the future work are presented in in Sect. 4.

2 The Adaptive Course in GALE

GALE is an extensible generic and general purpose adaptive hypermedia engine [12,13]. GALE is an evolution of the AHA! system [8]. AHA! was originally designed to serve a hypertext course taught through the Web [7]. It was later made more and more generic and extensible. As part of the EU FP7 project GRAPPLE it was completely redesigned and extended to become GALE. An adaptive course in GALE consists of concepts that are connected to pages. The pages contain information (or tests, exercises, assignments or anything else that can be represented by Web pages) and links to concepts. Links in GALE always refer to concepts instead of pages, and the link destination page may be adaptively selected by the adaptation engine. A page may also contain links to external pages on the Web, but we shall not consider these here. One approach used in GALE to adapt a course is to use different colors for the links of the pages. The author of the course defines rules to determine the conditions under which a presentation class is associated with a link. The default presentation classes are named “bad”, “good”, and “neutral”, and have the following meaning and presentation style:

1. The *bad* links point to non-recommended concepts, which means that according to the rules defined by the author, the student is expected to study something else—do some reading or perform some tests—before accessing these concepts. *Bad* links are colored in black and are not underlined, which implies that they are indistinguishable from the textual information of the page. So, *bad* links are hidden within the text, though they are fully operational and can be clicked on at any time.
2. The *good* links point to a recommended concept that the student has not yet visited after it became recommended. *Good* links are colored in blue.
3. The *neutral* links point to a recommended concept that the student has already visited after it became recommended. *Neutral* links are colored in purple.

Because of the choice of adaptation and colors the hypertext course looks like a standard (non-adaptive) website with pages and blue and purple links. In Fig. 1 we show an example screen-shot of this the hypertext course as served by GALE (and called upon from the Sakai learning management system).

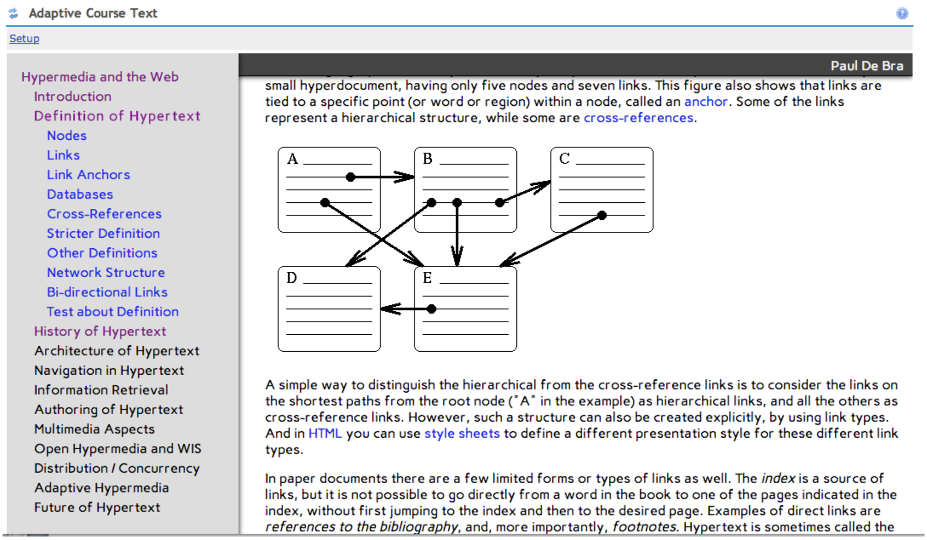


Fig. 1. Screenshot of a page from the hypertext course, showing also the navigation menu (left) and a header that gives access to settings and progress information.

The adaptive course evaluated in this paper was offered to bachelor students at the Eindhoven University of Technology. It is an adaptive course that introduces the basics concepts of hypermedia and the Web. This course has existed for a long time but the content was updated several times and the course as a whole was moved from AHA! to GALE. This has also allowed us to use the extensibility of GALE to perform more extensive logging, as needed for our evaluation.

At the beginning of the course (upon the very first access, and only on that access), the student gets a page with some explanation about the course and adaptation that is used. An important message to the student in this *intro* page is that in (almost) all the cases when the student request a page, the adaptation process is done, even when the student clicks on the *back* button of the browser. The message in this *intro* concept presents the following sentences:

The course pages are adapted each time you request them from the server. This should also happen when you use the "back" button of the browser, but some browsers (or browser configurations) may refuse to do so in which case you don't see the adaptation. There is nothing we can do about this.
...
Links in this course are not underlined (no, there is nothing wrong with your browser). Links appear mainly in three colors: ***blue***, ***purple*** *and* ***black***. *Since the main text is also black there may be links you don't see. These links are black because the system thinks these links are not (sufficiently) appropriate for you at the time you view the page containing them. Don't be alarmed, when you revisit the page later these links may become blue (or purple). The link colors are determined by the server, not your browser. You should not configure your browser to override the colors requested by the server.*

This message in the intro alerts the students about the importance of the link adaptation. The author suggests that students should not follow black links. The black links, as the course presents, "...are not (sufficiently) appropriate for you...". The author's intention is to tell the student to follow the structure of the course in a way that at some point they will have all the links colored in blue or purple.

In earlier analysis [6], we noticed that students followed roughly 6% of the links via black links, but we could not ask them why they did that, because we analyzed the logs long after the end of the course. For our new evaluation we implemented a plugin in GALE (you can find more detail about plugin implementations in GALE in [12,13]) that presents a multiple choice question whenever the student has followed 5 non-recommended links. We called this plugin *quiz*.

The students all started the course at the same time and they were aware that after three weeks there would be an exam about the course's content. Consequently, they had 21 days to learn about hypertext by studying the on-line course text. They were also aware that they would be allowed to consult and navigate through the course text during the exam. (In previous years students were free to take the course at their own pace and complete it with an assignment for which there was no deadline. This led to students postponing the work, concentrating on other courses with fixed deadlines and exams first.)

The main goal of this paper is to understand the influence of adaptation on the students learning in an adaptive course. One of our approaches to achieve the paper's goal was to investigate and try to understand why the students follow bad (non-recommended, black) links. We present our findings in Sect. 3. Another goal was to find a way of stopping the students from following bad links, or at least decrease the percentage of that. Our suggestions to course authors are presented in Sect. 4.

2.1 The Structure and a Few Rules of Adaptation of the Course

The course is composed of 175 concepts, including: 11 multiple choice tests, one about each major topic except the last one ("future"), and 1 final test (to prepare for the exam). From all the concepts, 12 of them are called the main concepts. We consider the main concepts to be the concepts that introduce a major topic in the course; the other concepts are sub-concepts of some main concept. Thinking of the course as a book, the main concepts would be the chapters with an introduction and the others would be the book's sections and pages. If we consider the adaptive course as a pure hypertext course, with no adaptation, the big difference between it and a book is that the student could find the same section in different chapters. Another difference is that the students do not follow a fixed page order, they follow embedded links in concepts in any way they like. In the adaptive case, the embedded links and their presentation class suggest how students should navigate through the course. Non-recommended links are indistinguishable from the running text and therefore unlikely to be followed by students. In the navigation menu however they are quite visible and rely on the (black) color to make their non-recommended status clear to the student.

The main concepts are listed in (and linked from) the course's welcome page. The welcome page is linked to a concept and it is presented every time a student logs in, unless the student tries to access a concept with a direct URL but he is not logged in (or his session is expired). In that case, after the student logs in the system redirects him to the concept he tried to access.

The students can set up the course to present the navigation menu shown in Fig. 1 or to hide that menu. The menu shows the main concepts and the sub-concepts of the currently active main concept. The menu links are also adapted like the links in the pages.

Initially, on the welcome page and in the menu, the system presents 3 out of the 12 main concepts as good links (recommended). The first 3 recommended main concepts are: *introduction*, *definition* and *history*. At the beginning of the course, there are 74 recommended concepts (good links), 1 visited concept (the Intro concept) and 100 non-recommended concepts. In the list of the 74 recommended concepts there are 3 tests related to the 3 (already) recommended main concepts. After the student performs these first 3 tests, 90 out of 100 non-recommended links become recommended (so the chances of students still accessing information through a non-recommended link become quite small after that point).

The welcome page also shows a sentence saying that the student has to perform all of the tests in order to be able to perform the final test. Indeed, to access the final test the student has to perform all the 11 tests in the course. It is important to remark that the final test is not just presented as a non-recommended link, it is blocked until the student performs the other tests.

The student can perform each test only once, except for the final test. At the final test, the system lets the student repeat it until he scores more than 90%.

The next section presents the evaluation data and analysis.

3 Qualitative and Quantitative Evaluation

The data considered for this paper consists of the access log of 46 bachelor students of the adaptive course offered in 2013 and a questionnaire answered by 28 of the same group of students. These were first year students in the programs of Web Science, Software Science and Psychology and Technology. The on-line course text was intended to prepare the students for the exam and for a later (group) assignment. The questionnaire was made available online to all the students, shortly before the exam (and left opens until a few days after the exam). The students were not required to complete the questionnaire and the system did not log their identity. The anonymous nature of the questionnaire should give students more "freedom" to answer the questions honestly. The course's access log contains 17,318 entries, where 609 entries are answers for the tests and 1,001 entries are accesses to the test concepts (the questions). So clearly students sometimes requested a (multiple-choice) test, realized they did not yet know the answers and went back to study some course pages before trying to complete the test. Note that the tests are generated randomly from a larger collection of questions, so when a student views a test but does not answer the questions the next visit to the test may show different questions. This is also true for the final test that can be repeated: each time the student will get (some) different questions.

We start our analysis by relating the test log with the exam grades (or to simplify, just grade). Students were told that the tests that are part of the adaptive course test are only there in preparation for the exam and do not contribute to the course grade. The grades that were obtained have an average of 5.2 out of a maximum of 10. (Grades were integers.) The highest grade was 7 and the lowest is 0 (only 1 student); the second lowest grade was 2 (only 2 students). 20 students obtained a grade above the average (i.e., a 6 or 7).

To understand Table 1 with the attempts and scores for the final test it is important to understand the operation of the course: only students who completed the tests embedded in each of the main concepts (chapters) were allowed (and had access) to the final test. Out of the 46 students only 12 performed all preliminary tests and gained access to the final test. Of these 12 only 8 students actually attempted to complete the final test. Students were told they should attempt to score at least 90% on that final test. The test could be repeated as often as desired (until the score of 90% was reached). The preliminary tests could be done only once and their score was not used for deciding on access to the final test and did not influence the adaptation. The final test was tried 319 times (by only 8 students). It is interesting that 5 out of 8 students who tried the final test got an exam grade above the average. We also note that all the students who got a score higher than 90% in the final test got a grade of 6.0

or 7.0. These findings suggest that if a student followed the course and tried to get a high score in the final test, he would be better prepared for the final exam than the "average" student. Although this is just anecdotal evidence (with few students taking and passing the final test) it is good to know because before the students started the on-line course they were told that the final test would be a good preparation for the exam. The anecdotal evidence "proves" us right.

Table 1. Number of time a student answered the final test, his highest score and the exam's grade

Student	# Tries	Highest Score	Exam Grade
Std1	1	27	7.0
Std2	1	27	7.0
Std3	3	40	3.0
Std4	34	100	7.0
Std5	47	47	3.0
Std6	49	93	7.0
Std7	51	60	4.0
Std8	133	93	6.0

The first three tests associated with the three recommended main concepts (introduction, definition and history) were performed by 40 students. Performing these tests turned the "advanced" concepts into recommended concepts, so that was a clear motivation for taking these tests. These 40 students thus clearly saw the link color changes after completing the first three tests. The questionnaire has the question: *Regarding LINK ANNOTATION, where the system presents links in different colors, perhaps with additional icons, which of the following statements applies?* For this question the answers was (in parentheses there is the number of students who chose that answer):

- I was not aware of link annotation. (1)
- I only notice that visited links became purple. (6)
- The link annotation was clearly intended to offer guidance through the course. (10)
- The link annotation was manly intended to avoid to going to some pages. (2)
- No answer. (9)

The answers show that 12 out of 19 students were influenced by the link annotation, since they recognized the importance of the link annotation to guide them through the course or to avoid them following links.

Even though students who performed the first 3 tests then got many more recommended links, they did perform some navigation through bad links before the advanced topics became recommended. The logs contain a total of 471 visits via bad links by 36 different students. To understand why the students would visit non-recommended concepts, the system presented a question to the students who accessed more than 5 concepts via bad links: *Why are you following black links (not suitable yet)?* The possible answers were:

1. I am curious to see what happens if I click on a black link.
2. I would like to explore the course a little bit before learning.
3. The system presented a lot of black links and I would like to know what it means.
4. I read many times about one concept and it is not suitable yet, but I would like to learn about it.
5. I do not know what a black link means.
6. Other (please specify).

15 students visited less than 5 concepts via bad links and did not get this question. We only wanted to question the “repeat offenders”.

Table 2 presents the summary of the students’ answers. It is interesting to note that 13 out of 18 answers were given on the day of the exam, including 3 of them during the exam, showing that students studied the course text just in preparation for the exam, not really considering that the most important aspect of the course was that it prepared them for the later assignment. One student, called *Std1*, chose the *other* answer during the exam and he wrote: “Trying to find an answer to a question”. Std1 had already answered this question a few hours before, and at that time he chose the answer 2. This student reset his profile after the first “exploration phase”, which means that the system deleted everything about what he did from his user model, including the performed tests and visited status of concepts. (The log however remained unaffected by a reset.) We can also estimate the time students spent reading pages, by considering the time difference between two log entries for that student. Std1 showed an average reading time of roughly 9 seconds per page. Clearly in the exploration he did not spend enough time on the pages to actually read and study the text. The system adaptation influenced Std1’s steps in the course in spite of the fact that he was not concerned with learning the material and was only exploring the course. The adaptation in the course only depends on access, not on reading time. Because of the profile reset this student received the question about navigating through black links twice. And it turns out he gave a different answer the second time.

Table 2 also summarizes how we could groups students who do not follow the system recommendation:

Table 2. Summary of the answers about why students follow bad links

Answer	Before Exam Day	Exam Day	**Total**
1	2	2	**4**
2	2	6	**8**
3	1	0	**1**
4	0	2	**2**
5	0	1	**1**
6	0	2	**2**
Total	**5**	**13**	**18**

- students who have curiosity about what would happen if they do not follow the adaptive structure. It is represented by students who answered 1, 3 and 4;
- students exploring the course: students who want to explore the course before they start learning. It is represented by students who answered 2;
- late students: students who start studying the course when it is (too) late. Three of them were still following bad links during the exam.

As it appears that many students started studying for the exam quite late it is interesting to look specifically at the log for the day of the exam.

On that day the bad links access log shows 278 accesses (including during the exam), which is roughly 60% of the all bad link accesses. During the exam there are 168 accesses via bad links, representing roughly 36% of all bad links accesses. These 168 accesses are concentrated in 9 different students, where 3 of them accessed only 1 bad link each and the others have an average of 27.5 (4 students above the average). Teachers often state without proof that some students go to an open book exam without any preparation, thinking they can look everything up as needed during the exam. Our log proves that some students indeed did not prepare for the exam and tried to search for the answers during the exam, and thus often going through bad links. Before the exam day the bad link accesses are mostly concentrated in the log for 5 students who performed 120 of these accesses, representing roughly 25% of all bad links accesses. These students are in the category of student curiosity or student exploring the course according to their answer on the question about why they followed bad links.

In the questionnaire, the students answered the question: *Regarding LINK HIDING, where the main text is black, the black link is hidden in the main text but you can still click on it. Have you tried clicking on these links?* For this question, we have similar results to the one presented in Table 2: 67% of the students who admitted to have followed black links talked about the curiosity; 25% admitted to explore the course before learning; and the 8% said that they did not remember if they followed black links.

Another point to corroborate the idea of the students' curiosity is the fact that the first two most accessed concepts via bad links are the first two of the main non-recommended concepts appearing in the list of topics on the welcome page. (These are "architecture" and "navigation".) The referrer of a visited concept indicates the concept the student came from. So we could also analyze whether other non-recommended concepts were visited directly from these concepts or whether they were accessed through cross-reference links from other main concepts. So we looked at the all pages from which bad links were followed. Again, the navigation and the architecture concepts appear in the top four rank, losing the first and second positions only to the welcome concept that lists all the main topics and the *TO DO* list (a page containing a list of all the concepts that were not visited by the student before.), two links that are always recommended. This indicates that when a student followed a bad link to a main topic he was not considered ready for, he did not stop at the first page of the non-recommended topic but kept following bad links on that topic. When asked about this behavior the answer to the bad link question was "by curiosity".

The use of link colors in the adaptive course text is described as *link hiding* because black link anchors appearing in a paragraph of black text in fact causes the link anchor to be hidden. However, the black links also appear more clearly visible in three places: on the welcome page that shows a list of links to the main concepts (Fig. 2 show a screenshot of this page), in the optional navigation menu of the course (also shown in Fig. 2), also showing bad links and in the TO DO list (of all unvisited concepts). In these places the differences in color should be referred to more correctly as *link annotation* instead of *link hiding*. For example, Fig. 2 illustrates the presence of bad links. In the menu on the left, the items from "Architecture of Hypertext" to "Future of Hypertext" are web links in black color. Similarly, in the welcome concept (right) shown in Fig. 2, the list of topics from "The architecture of hypertext systems" to "Assignment for this course" are web links in black color that can be clicked by the student. We can thus distinguish between bad links followed through annotated links versus bad links followed through hidden links. It turns out that 87% of all bad link accesses are through the three mentioned places that use *link annotation*. When we use the same three sources (welcome, menu and TO DO list) to check the accesses via good links, the percentage falls down to 57%. The accesses that remain thus indicate that the students follow many links from the course pages but few bad links in the course pages. Link hiding is thus effective in keeping students away from non-recommended topics but link annotation is not. Or in other words, students look for good links when they are navigating through the course pages but consider all the links that appear in menu-like structures.

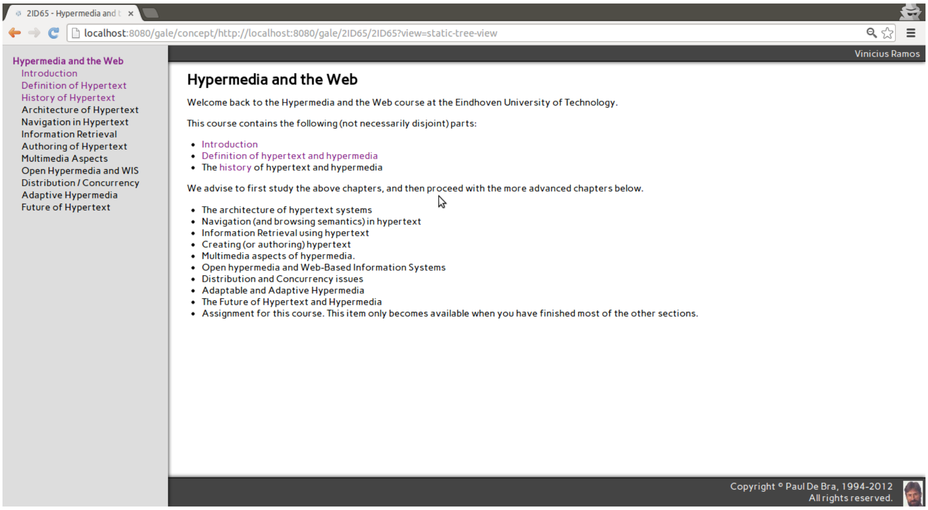


Fig. 2. Screenshot of the *welcome* page from the hypertext course, showing also *bad* and *neutral* links in the menu (left) and in the main view (right)

We asked the students: *Regarding adaptation, what was your user experience?* The students answers seem to contradict the logs: 8 out of 28 students (28%) marked the option: "I hardly noticed it", while 9 others marked the option "I found it patronizing", and only one "found it helpful". 10 students did not give an answer at all. It is important to recall that roughly 90% of the students completed the first 3 tests of the course. Consequently, they got 90 new recommended concepts, of which 9 are presented in the menu and in the welcome page. For this reason, it is easy to see that the adaptation is working. Yet 28% of the students said they "hardly noticed it". Apparently link hiding and link annotation as used in this course (having links become recommended but never making a recommended link become non-recommended) is a fairly unobtrusive adaptation technique.

Considering that 51.5% of the log entries correspond to the last day of learning and 26.5% of the whole log belongs to the day of the exam (including during the exam), it appears that the students were much more concerned with quickly preparing themselves for the exam and were not too concerned with a comfortable learning experience for which the adaptation is intended. The log entries accumulated until two days before the exam represent 23% of the whole log, and the bad links entries on these days represent only 15% of the total number of bad links followed. Students who started learning in the beginning of the course period followed the course's link structure and adaptation better than of the students who started learning a few days (or even hours) before the exam. Clearly, when rushing towards the exam students are not open for advice that tells them they are not yet ready for certain topics and should study something else first. They want to take it all in, as quickly as possible and in any order.

When students were asked to talk about their feelings about the adaptation in the course, two of them talked about the menu that could be enabled or disabled. They would like the menu to be enabled by default, whereas in the course it was disabled by default. They see an advantage of following the menu while following links, instead of relying on links that appear within the pages. Another student said that *It's not my cup of tea. I see the advantages, but I prefer an old-fashioned paper book.* Two students also suggested the possibility of having a back link/button within each page of the course, to be used instead of the browser's back button, so as to allow them to keep navigating through the pages' content.

In the next section, we discuss the results presented here and the suggestions made by the students in the questionnaire. We also present our suggestions to the author of the course and adaptive course authors in general.

4 Discuss and Future Work

In this paper we presented the evaluation of a distance learning course implemented in GALE. The focus of this case study has been to analyze qualitatively and quantitatively whether the adaptation mechanism influences the students' learning and navigation behavior. We approached these topics in two ways: verifying the access log of the course, comparing it with the exam grades and analyzing a questionnaire answered by the students. We used the extensibility of GALE to add logging of the link annotation/hiding states and to automatically ask students about their use of bad (non-recommended) links.

The main analysis of the access log relates to the visits of concepts via the so-called bad links (that point to non-recommended concepts) and whether the presence of these links affected the student's behavior. We found that students who start studying late in the course period tend to follow more bad links than students who use the entire period. We also noticed in the log and questionnaire analysis that students often follow bad links because they are curious about the linked non-recommended concept or they are exploring the course before they start studying more in depth. We have observed that students navigating via bad links continue through bad links for a few more links. When students navigate through bad links 87% of the clicks came from the welcome page, the menu view links or the TO DO list of concepts. On the one hand this suggests that they are curious or exploring the course before they start learning as confirmed by them in the questionnaire and by answering a specific question about following bad links during their navigation. On the other hand it also shows that the *link annotation* used on these three pages has less effect on the students' navigation than the *link hiding* that is used for links appearing in the running text of the pages.

An important remark about the access log is related to the fact that 51.5% of the entries occurred on the day of the exam (including during the exam), and 26.5% of the entries occurred on the day before the exam. There are more entries during the exam (4,495 entries) than during the whole period before the day before the exam (3,976 entries). This suggests that, independently of the exam (which the adaptive course was supposed to help the student to study for), the students need to be stimulated to start studying right from the beginning of the course term. The absence of classes or other contact with a tutor during the course period leads to students postponing their study activity until right before the exam.

The questionnaire gave us a good insight in the needs of the students, because they had a few suggestions to improve the quality of the course and its navigation. Students suggested to have a navigation menu presented at all times (by default). We made this optional because the course was on the topic of hypermedia and we wanted to stimulate navigation by following links in the pages. But for other course topics it is certainly worthwhile to offer a navigation menu by default. Some students remarked that the course structure (i.e., the navigation support including a menu) should be the way they are used to study, more like a book. The new paradigm of navigating through links in pages makes the navigation, as a student wrote, "quite difficult and annoying."

Students also want to know how far along they are in the course. Such a "count down" counter exists in the course but it is not permanently displayed. Instead of a menu a progress bar could also be used. As the deadline (the exam) was approaching it is clear that students lose their patience in following the advice to study some concepts before some other concepts. They want to click through the whole course quickly.

For future work we decided to implement a analysis tool to allow the authors of the courses to make their own analysis, in real time, while students are using the course text. We also plan to carry out another case study of the adaptive course with master students who will implement their own extensions to GALE in order to run other analysis experiments.

References

1. Brusilovsky, P.: Methods and techniques of adaptive hypermedia. User Modeling and User Adapted Interaction 6(2–3), 87–129 (1996)
2. Brusilovsky, P.: Adaptive hypermedia. User Modeling and User Adapted Interaction 11(1-2), 87–110 (2001)
3. Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.): Adaptive Web 2007. LNCS, vol. 4321. Springer, Heidelberg (2007)
4. Brusilovsky, P., Schwarz, E., Weber, G.: Elm-art: An intelligent tutoring system on world wide web. In: Frasson, C., Gauthier, G., Lesgold, A. (eds.) ITS 1996. LNCS, vol. 1086, pp. 261–269. Springer, Heidelberg (1996)
5. Brusilovsky, P., Sosnovsky, S., Yudelson, M.: Addictive links: The motivational value of adaptive link annotation. New Review of Hypermedia and Multimedia 15(1), 97–118 (2009)
6. Culmant-Ramos, V.F., De Bra, P.: The influence of adaptation on hypertext structures and navigation. In: Proceedings of the 21st ACM Conference on Hypertext and Hypermedia, HT 2010, pp. 77–82. ACM, New York (2010), http://doi.acm.org/10.1145/1810617.1810632
7. De Bra, P.: Teaching hypertext and hypermedia through the Web. Journal of Universal Computer Science 2(12), 797–804 (1996)
8. De Bra, P., Smits, D., Stash, N.: The design of AHA! In: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, HYPERTEXT 2006, pp. 171–195. ACM, New (2006), http://doi.acm.org/10.1145/1149941.1149942
9. Gena, C.: Methods and techniques for the evaluation of user-adaptive systems. Knowl. Eng. Rev. 20(1), 1–37 (2005), http://dx.doi.org/10.1017/S0269888905000299
10. Knutov, E., De Bra, P., Pechenizkiy, M.: Ah 12 years later: A comprehensive survey of adaptive hypermedia methods and techniques. New Review of Hypermedia and Multimedia 15(1), 5–38 (2009)
11. Mulwa, C., Lawless, S., Sharp, M., Arnedillo-Sanchez, I., Wade, V.: Adaptive educational hypermedia systems in technology enhanced learning: A literature review. In: Proceedings of the 2010 ACM Conference on Information Technology Education, SIGITE 2010, pp. 73–84. ACM, New York (2010), http://doi.acm.org/10.1145/1867651.1867672
12. Smits, D.: Towards a Generic Distributed Adaptive Hypermedia Environment. Ph.D. thesis, Eindhoven University of Technology (2012)
13. Smits, D., De Bra, P.: Gale: A highly extensible adaptive hypermedia engine. In: Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia, HT 2011, pp. 63–72. ACM, New York (2011), http://doi.acm.org/10.1145/1995966.1995978

Evaluating the Use of Open Badges in an Open Learning Environment

Jose Luis Santos[1], Sven Charleer[1], Gonzalo Parra[1], Joris Klerkx[1], Erik Duval[1], and Katrien Verbert[2]

[1] Dept. of Computer Science, KU Leuven
Leuven, Belgium
{JoseLuis.Santos,Sven.Charleer,Gonzalo.Parra, Joris.Klerkx,Erik.Duval}@cs.kuleuven.be
[2] Dept. of Computer Science, Eindhoven University of Technology
Eindhoven, The Netherlands
k.verbert@tue.nl

Abstract. This paper reports on our ongoing research around learning analytics. We focus on how learning analytics can be used to increase student motivation and the use of badges as a way to aggregate learning activity being a representation of their goals and progress along the course. The context of this work is an open learning environment, based on wikis, blogs, twitter, an activity stream mash-up and an open badges system. Our evaluation analyses perceived usefulness and usability of the system, as well as the impact on student motivation. Our results indicate that badges are useful to motivate students while activity streams have the potential to activate students.

1 Introduction

Motivation, while often overlooked, is perhaps the single most important factor in learning [1]. Also in our earlier work [2], students reported motivation as one of their main learning issues. New strategies to raise learner motivation have been discussed in recent years [3,4].

Some research identifies the factors that influence motivation in order to plan interventions [5]. Others study how technology such as using augmented reality devices [6] and learning dashboards [7] can impact on the learner motivation. Whilst nobody has found a magic recipe, we try to figure out what strategy works better in our context.

We believe that social support and collaboration among students and active teacher support may help to overcome and regulate problems in motivation [1]. In order to promote such collaboration in our courses (see section 3.1), we follow an 'open learning' approach. Students use Twitter (with a course hashtag), wikis, blogs and other web 2.0 tools such as Toggl [2] and TiNYARM [8] (i) to report on and (ii) communicate about their work with each other and the outside world in a 'community of practice' kind of way [9]. We focus on the use of learning dashboards that visualise data about learners in an open way, therefore students

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 314–327, 2013.

can check their own and peers' information [2]. We design, build and evaluate the dashboards in order to measure their potential to support awareness and (self-)reflection and to motivate students to participate actively [2]. Whilst several of our case studies showed that these kind of dashboards can successfully address a variety of learning issues, one thing they did not, is to increase learner motivation [10,2]. This paper focuses on the next version of our dashboards of which we believed they could increase learner motivation.

The learning dashboard described in this paper merges three approaches:

- **Badges.** Exciting new research is adopting badge systems to certify skills and abilities [11] such as Mozilla Open Badges that allows students to share their achievements/badges on different social networks such as blogs, Linkedin and Facebook [12]. The use of badges have been proven to positively increase learner motivation [13]. In addition, badges are a representation of how students carried out the activities which in turn can affect positively the learner motivation [14].
- **Social visualisations.** The use of social visualisations may increase learner motivation [7].
- **Activity Stream**. Activity streams increase user awareness [15] by aggregating the activity from the different systems that are being used in our course: posts, comments, tweets and badges.

We carried out a detailed online 5-point Likert scales questionnaire and a second evaluation with open questions four weeks into the course. 21 out of 26 students replied to the survey. The questionnaire assesses perceived usefulness and usability of the different components of the system.

The remainder of this text is structured as follows: section 2 describes related work. Section 3 describes the architecture of the Open Learning Analytics system and the badges used in this experiment. Section 4 explains the evaluation settings and the obtained results. Conclusions and future work are presented in Section 5.

2 Related Work

Dashboards can be considered as a subset of 'personal informatics' applications[16]. These applications typically encourage positive change of behaviour or provide new insights from data[17,18].

Results of our analysis of learning dashboards [17] suggests that dashboards can have an impact on the behaviour of students if they go through four stages:

- The first stage is awareness. This level concerns with just data. Such data can be visualised as activity streams, tabular overviews, linecharts or other visualisations.
- The second stage is reflection. It concerns users coming up with questions and assessing how useful and relevant these are.
- The third stage is sense-making. This level concerns users answering such questions.

- The final stage is impact. This level concerns new meaning or change behaviour.

Dashboards such as Moodle dashboard [19] visualise the overview of the activity in a table. In addition to the table overview, StepUp! [2] visualises the activity using also spark lines and bar charts to zoom in the activity .Other dashboards visualise the data using line charts such as GLASS [20] and others are complemented with parallel coordinates and word clouds such as the Student Activity Meter [10]. Class-on [21], for instance, visually represent the class and the status of the student with different colours for each seat.

The dashboard described in this paper visualises the achievement of Open Badges along the course using line charts. Open Badges [12] are used as a mean to aggregate data and representation of achievements. In addition, Badges are also completely public enabling factors such as social influence and the perceived degree of recognition [22]. This dashboard is supported also by an activity stream mashups providing a quick and easily accessible overview of a group's work. Different activity streams enable users to become peripherally aware of contributions[15]. Some activity stream dashboards visualise this information by aggregating information from different social networks in a single stream [23]. The activity stream presented in this paper also aggregates information when students are awarded with badges.

Depending on how the data is visualised, the dashboard will trigger different questions and consequently different conclusions [24]. Some dashboards are designed to detect isolated students [25] or students who do not work well [26]. Other dashboards are used to increase awareness of class activity [2] and resource use [10] or to increase motivation of students [7]. Our learning dashboards is designed specifically for motivational purposes. Badges are set up in order to represent the way that activities are carried out. Displaying this information should motivate the students [14].

3 Open Learning Analytics System

3.1 Introduction

The system described in this section has been designed based on the previous statements and deployed in a course on user interfaces where 26 engineering students work in teams on projects related to Human-Computer Interaction. This year they focus on creating a "recommendation application". A large part of the course activity focuses on the design, implementation and evaluation of a working prototype, starting from a paper prototype to a working digital prototype. Each group sets up a blog on which they regularly post updates on thoughts, ideas and progression. They also tweet using a course hashtag (#chikul13) and comment on each other's blogs. All blogs are deployed on the Wordpress[1] platform. This allows us to easily track student activity through the use of RSS

[1] http://developer.wordpress.com/docs/api/

feeds and the availability of their usernames for all comments. Interaction with teacher and assistants outside the classroom also happens through Twitter and comments on their blogs. Using these platforms, we create an open course that is accessible by everyone, promoting interaction with external parties and giving the tasks a value also outside of the course.

3.2 Architecture

Our system (see Figure 1) consists of trackers, the Activity Stream, the Badge Rewarding System, Navi and the Data Store.

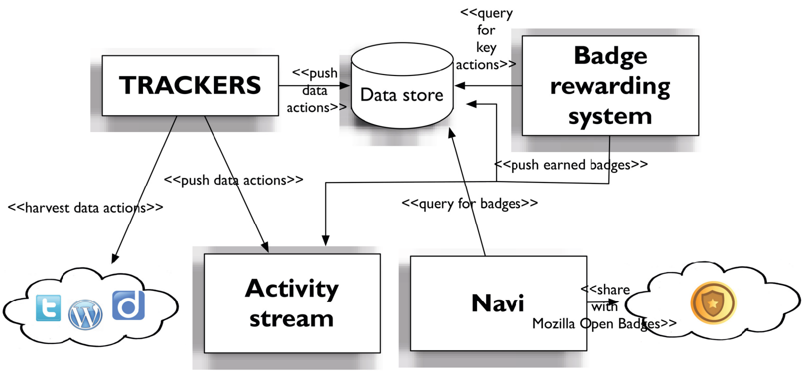


Fig. 1. System architecture

- The trackers aggregate the RSS feeds of the students' Wordpress blog posts and comments and the Twitter[2] activity related to the course hashtag.
- The activity stream (see section 3.3) visualises these streams of information and relies on the TiNYARM software. The Badge Rewarding System contains and applies the rules that define who will be awarded a badge (explained in Section 3.4).
- Navi (see section 3.4) visualises the badge information and follows the Mozilla Open Badges Infrastructure[12], creating interoperable badges that the students can share across different social networks.
- The Data Store is the glue between these applications as it stores the information streams and exposes them through a RESTful web service.

These applications work independently and run on Google App Engine[3]. The Data Store is hosted on a server running Apache Tomcat[4] and PostgreSQL[5].

[2] https://dev.twitter.com/docs/api
[3] https://appengine.google.com
[4] http://tomcat.apache.org/
[5] http://www.postgresql.org/

3.3 Activity Stream

Following the activities in the course with a fragmented set of tools can easily become a time consuming and disorienting task. With this in mind, we have deployed a web application to provide a quick and easy way to get an overview of the course activities. The Activity Stream (AS) receives output from the different applications and presents it to the students as an social awareness stream[27]. The AS (see Figure 2) shows the tweets, blog posts, comments, bookmarks, and badges of each student. These activities are composed of 4 elements: authorship, action, artefact, and date; providing answers to the questions of: Who is doing what?, What are they doing?, and When did that event happen?. The students therefor get a general overview of the course, filtered streams of activities based on the applications used, and an activity profile of their classmates. These views allow students to quickly get a complete summary of what is happening in the course. The different streams are also available through RSS feeds to provide another way to consume this information.

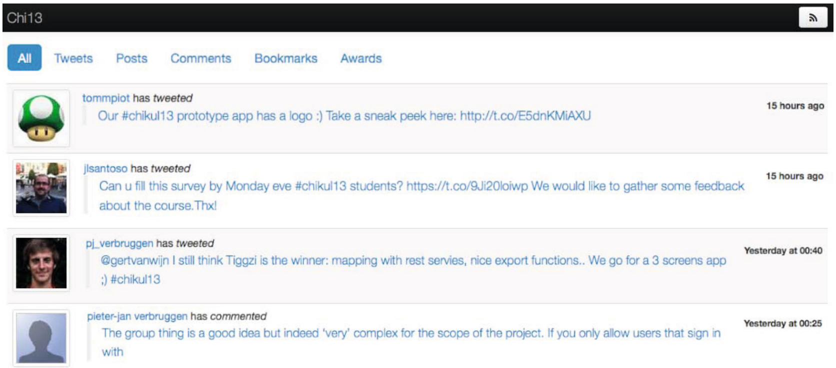


Fig. 2. Activity Stream main screen. It can be visited at http://chi13course.appspot.com/

The AS is an aggregator or mash-up of activities, same as any RSS reader or TweetDeck[6]. The course activities are pushed to the application by different trackers following the Activity Stream specification [7] and structured as: Actor |verb |(Object). For example, jlsantoso |has tweeted |'Can u fill this survey...'. Other examples are available in Figure 2.

3.4 Navi

During the course, students are rewarded for their activity and achievements through the means of badges. These badges can shed light on their progression

[6] http://tweetdeck.com/
[7] http://activitystrea.ms/

and that of the class. To design these badges, we start from the idea of what we want to promote, namely: activity, quality and results.

Evident choices for activity are the amount of entries posted on Twitter and Wordpress, resulting in, for example, badges rewarding 5, 10 or 15 tweets posted. Quality can be indirectly linked to the external activity on a student's activity. For instance, a blog post with many comments by external parties or "retweets" in Twitter can be a good indicator, hence we design badges rewarding the acquiring of 5, 10 or 15 comments on one's blogpost. Results are related to milestones so students can be rewarded for completing certain activities such as finishing an assignment. Thus, we award badges for the completion of a development iteration in the course.

To keep the students actively engaged during the entire duration of the course, most badges are awarded biweekly. There are badges that are awarded instantly (i.e. a student that tweets 5 times receives a badge), other badges are awarded at the end of the 2 week period (i.e. the most active student in class). A few badges can only be achieved once in the course (i.e. development iteration 1 completed). To reward team effort, some badges can only be earned as a team. Keeping track of and commenting on every blog post is more feasible as a team and is rewarded with an appropriate team badge. While generally badges represent positive activities, we designed a couple of negative badges that students should not wish to receive, such as the badge for 'no comment' activity.

To visualise this information, the dashboard application Navi was developed giving the students an overview of the badges that have been awarded by the badge system. This application has been developed in Java and Javascript, using D3.js[8] and is deployed on Google App Engine.

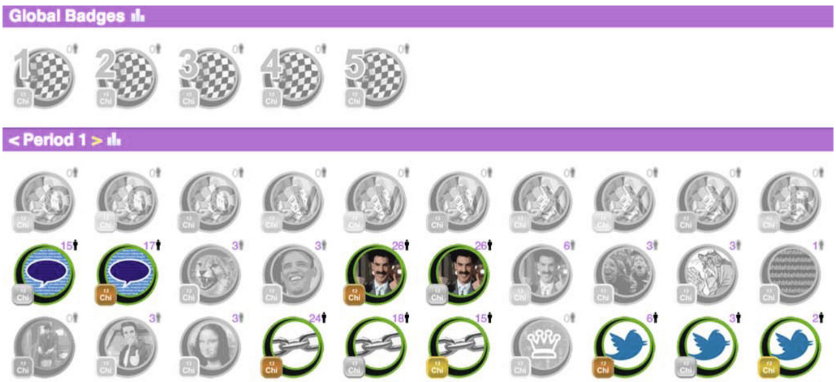


Fig. 3. Navi Personal Badge Dashboard: achieved badges are coloured. The number next to the badges indicates the class's progression for the badge. It can be visited at http://navi-hci.appspot.com/

[8] http://d3js.org/

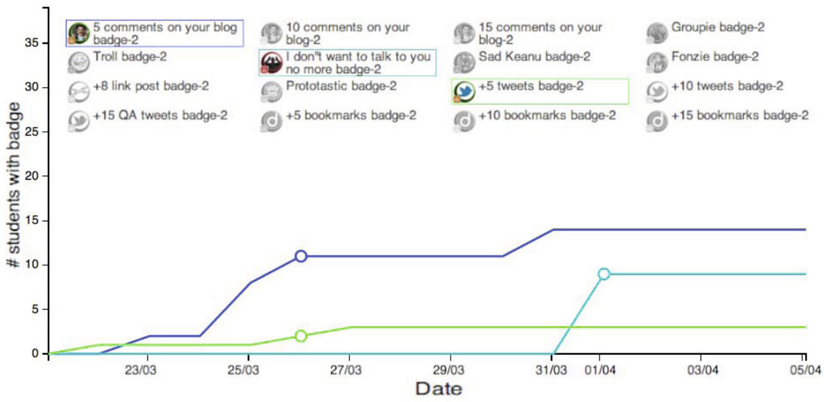


Fig. 4. Navi Class Progression: Each line represents class progression for a specific badge. The circle indicates when the student achieved the specific badge.

At the start of the course, we present the students a list of all achievable badges. Every definition of a badge is set up as follows:

- a badge icon with an easily identifiable image related to the meaning of the badge
- a color coding for categorizing the badge by type
- a bronze/silver/gold medal for badges indicates different steps of achievements
- a textual description on how this badge can be achieved.

This list is available in Navi and helps students understand what activities and results are required to complete the course.

Figure 3 illustrates Navi's Personal Dashboard which contains the achieved and remaining badges of one student. To support a better group awareness, Navi displays the total number of awarded badges next to each badge. Hovering the mouse over a badge displays the names of the students who have acquired this badge. Students can also compare their own progress with that of their peers.

Drilling down on badges in Navi's Personal Dashboard, the student obtains Navi's Class Progression view, a visualisation of the awarded badges over time. An example of this view is shown in figure 4 for student A where the X-axis represents time and the Y-axis the number of students that have been rewarded the badge. Each line thus represents the class progression for a specific badge. The circle indicates when student A was awarded this badge. With these views, Navi invites the students to reflect on their individual progress by giving them a tool to be aware of their activities, but also of that of their fellow students.

Navi's Personal Dashboards are also available to teachers, providing them with an immediate overview of the class progression. The dashboards can help teachers figure out what activities the class or even specific students are struggling with.

4 Evaluation

4.1 Evaluation Setup, Data and Demographics

We carried out a detailed online 5-point Likert scales questionnaire four weeks into the course. This survey was not anonymised: we informed the users that their name was required to find possible correlations with their activity within the community for research purposes. In the questionnaire, we analysed four aspects, in order to obtain a broad view:

- Frequency of use. How frequently students visit the applications and why (see marker 1 at figure 5, the colors are defined in the legends below the boxplots). The results are discussed in Section 4.2.
- usability of the activity stream and Navi (see marker 2 at figure 5). In addition we use the SUS questionnaire[28]. SUS scores and results are discussed in Section 4.3.
- usefulness and how they perceive that Navi and the activity stream are helping them (see marker 3 at figure 5). The results are discussed in Section 4.4.
- perceived usefulness of the badges (The results are not visualised but they are discussed in Section 4.5).

A week later, we performed a second evaluation to gather additional information regarding usefulness and usability that was not clear in the first survey. This second evaluation relies on open questions and is anonymous.

Furthermore, we track the use of the wiki, the activity stream and Navi with Google Analytics[9]. This data together with the individual activity within the community (comments and tweets) are visualised in figure 6. The linechart helps us to find usage patterns.

These two additional sources are referenced along the sections 4.2, 4.3 and 4.4 in order to enrich the discussion.

21 students out of 26 replied to the online survey. They are between 20 and 25 years old. 20 out of 21 were males and only one female student replied to the survey. All participants are master students. 24 out of the 26 students participated in the second survey.

4.2 Frequency of Use

Google Analytics. The activity stream, Navi and wiki are all actively used (see Figure 6). Navi is the most actively used application based on figure 6. Moreover, the activity stream and wiki page views are aligned. Activity peaks correspond to the face-to-face sessions with the teachers.

Information about the Course. The survey results show that our students prefer the wiki, the url of the blogs and following the twitter hashtag for getting

[9] http://www.google.com/analytics/

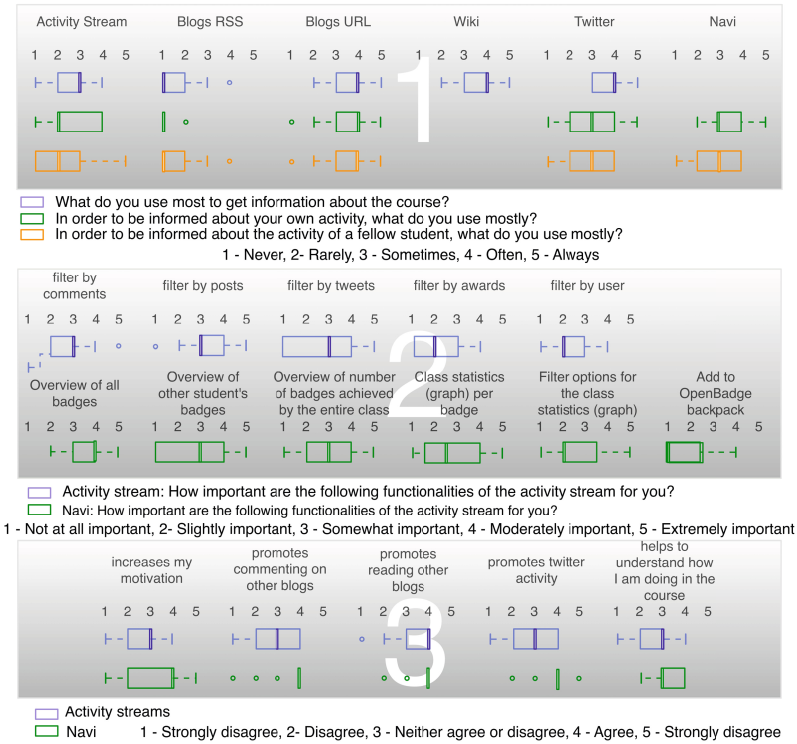


Fig. 5. Boxplots of the results gathered from the survey

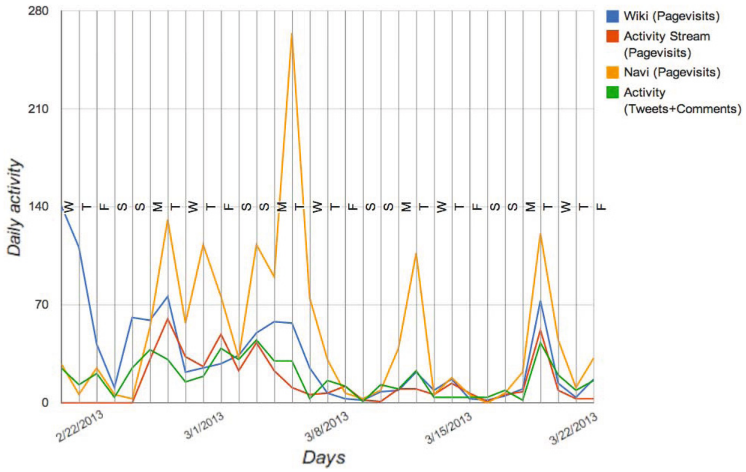


Fig. 6. Wiki, Activity Stream, Navi Google Analytics data and the student activity (tweets+comments). The letters correspond to the first letter of the week day.

information about the course (see marker 1 at figure 5: all of these have an average Likert score mean of 4). The activity stream is visited less (Likert score mean is 3). However it is used more than the RSS feeds which are almost not consumed (Likert score mean is 1). Based on Figure 6, we expected better results regarding the activity stream that was designed as a central point of information. More insights from the second evaluation from a usefulness perception are discussed in section 4.4.

Gaining Insight about their Own Activity and Peers Activity. Figure 5 (second and third row at marker 1) point out that Navi and the activity stream were rarely used (Likert score mean are 2 and 3) in order to gain some insight about their own or fellow students activity. Nevertheless, we highlight that the student distribution is slightly more positive when it concerns to individual activity. Based on Figure 6, we expected better results regarding the activity stream and Navi because both tools have a profile view. Whilst these results may point out that students do not perceive both tools for reflection purposes, the second evaluation provided more insights about perceived usefulness of the tools and they are further discussed in section 4.4.

4.3 Usability

SUS Scores. We analysed usability through SUS scores[28]: the results are a total score of 70 and 65 for the activity stream and the badge dashboards respectively. These results are below what is considered good usability.

Functionalities of the Activity Stream. First row in marker 2 at figure 5 shows that all the activity streams are considered quite neutral.

Regarding the activity stream, 15 out of 26 students highlighted that this provides an easy overview of all activity and avoids that they have to go through all the blogs. On the other hand, 10 out of 26 students reported an information overload problem with expressions such as *'Sometimes feels a bit cluttered'*. However, four of them state that the filters were useful to avoid this problem. In addition, 3 students proposed additional filters such as filtering per groups of students and filtering per time periods of the course.

15 students highlighted the main functionality of the activity stream compared with only 10 in the case of the dashboard. These results are reflected neither in the survey results or the use of the activity stream. Notifications triggered students to visit Navi, in the same way, this extra feature could improve the perceived usefulness of the activity stream. Moreover, we need to further research alternatives to deal with the information overload issue reported by the students and additional features to increase the perceived usefulness results.

Functionalities of Navi. Second row in marker 2 at figure 5 shows that 'Overview of all badges' is considered important (Likert score mean of 4), the rest of functionalities are quite neutral (Likert scores mean of 2 or 3).

Regarding Navi, 10 out of the 26 students who participated in the open ended feedback of the second evaluation highlighted the personal overview of the badges. 8 students also highlighted comparison of the earned badges with their peers. On the other hand, 3 students reported some problems regarding the graph functionalities. They reported that it is difficult to draw conclusions from the visualisations. 2 students reported that they only accessed the badge dashboard when they got a notification email that they earned a badge. 2 other students report that there are too many badges and that this makes it difficult to be fully aware of the ones they can earn. 9 students reported that they did not know the functionality of sharing their badges through Mozilla Open Badges.

Based on the quantity and quality results, we conclude that the 'Overview of all badges' is the most important functionality, although the ability to compare their achievements with their peers is also used. Addressing the graph usability issues reported by the students, students may use Navi more frequently in order to be informed about their own activity and their peers (section 4.2). Moreover the badges notification trigger some of them to visit Navi. This notification system can also be used to reinforce the use of the graphs or other new features in Navi and the activity stream.

4.4 Perceived Usefulness

Motivation. Whereas first column in marker 3 at figure 5 shows that the students consider badges motivating (Likert scale average of 4), they do not have same opinion about the activity stream (Likert scale mean of 3). This results are somehow logic since the activity stream is designed to increase the learner awareness and not motivation. However, it is also useful to see clearly the different scores between both applications in order to see how the perception changes regarding a tool designed to motivate and the other that was not designed with this goal.

Motivating Activity. Columns 2, 3 and 4 in marker 3 in figure 5 show that students believe that badges motivate reading other blogs, writing comments and tweets, whereas the activity stream only promotes reading other blogs (Likert scale mean of 4). However, we also know from earlier research that the activity stream may help them to decide what to read next [8].

How they are Doing in the Course. Results are inconclusive regarding either the activity stream or Navi helps them to understand how they are doing in the course, although the student distribution is slightly more positive in Navi's boxplot. We expected better results since badges are designed in order to indicate their progress. Further research is required in this area. Badges are further discussed in section 4.5.

Earlier Research. In earlier research[2], we evaluated StepUp!, a dashboard that displays the raw data of the community in a big table overview with the same questions discussed in this section. The obtained results were similar to the activity stream regarding motivation but similar to Navi regarding the other questions.

All three apps require tracking of student activity, they make the data public and they get the same result regarding promoting reading of blogs. Therefore increasing the student awareness of the activity within the community may promote the curiosity of the students to read what others are doing.

Whereas Navi and StepUp! aggregate the activity per actions such as comments and tweets in badges or simple numbers, the activity stream displays uniquely the individual actions. Therefore, the aggregation of the data may promote actions such as commenting or tweeting.

Navi uses badges in order to give the message to the user, whilst StepUp! relies the whole cognitive effort on the users who have to process the information and draw conclusions by themselves. Badges may provide this motivational element to the students.

Open Feedback in Second Evaluation. In the second evaluation, 4 out of 26 students also reported that the badge system was motivating and boosts extra effort for the course, whereas one student wrote the opposite. This kind of statement supports our previous analysis regarding motivation.

Regarding the activity stream, 3 out of 26 students highlighted that the centralisation of all the information was a nice feature for the course. 5 other students also reported that it was a nice way *'to catch up with what others are doing'* . This feedback also supports the conclusions of section 4.3. However we should explore the implementation of new features in order to improve the perceived usefulness such as a notification system as Navi has.

4.5 Perceived Usefulness of the Badges

We asked students to score 18 badges following a likert scale from not at all important to extremely important.

Summarising the results, they surprisingly consider among the most important badges those badges which have a negative connotation. The rest of the badges are ranked as we could expect. The most important ones are those who promote to achieve learning outcomes, afterwards those who want to improve the quality and finally those who want to promote quantity. Moreover, they consider group badges more important than individual badges, except for the negative ones that are in the top.

In the second evaluation, 2 students reported that there were many badges. However, we expect that badges increase their curiosity in order to explore them and to increase their learning motivation. Therefore, we do not consider it necessary to reduce the number of badges as the system is motivating them.

5 Conclusion and Future Work

This paper reports on an experiment where we deployed a badge reward system in order to increase learner motivation and activity stream to raise learner awareness about contributions within the learning community.

Comparing the results of three different tools such as Navi, StepUp! and the activity stream, we conclude that the mere fact of tracking student activity and making it public through student activity such as activity streams, tabular and badge overviews may motivate activity such as reading. Social visualisations such as StepUp! and Navi may trigger more explicit activity such as commenting and tweeting. Moreover, badges motivates our students.

Whereas students reported that they did not really use the graph visualisations, they reported that they used Navi to compare their badges with others in the second evaluation. Such comparison may trigger some reflection process. In order to simplify this comparison process, we are going to explore others visualisations and the use of tabletops and ambient displays in a classroom environment to enhance such social comparison enabling several people to interact with the visualisations at the same time and promoting social discourse.

Based on section 4.3, students did not add badges to their Mozilla Open Badges Backpack. Further research is required in order to gain some insight about what kind of badges students would like to share through such system.

Students also reported that the activity stream was a useful central point of information, however they also reported that they used it rarely. We have seen that notifications triggered visits on Navi. Email notifications may raise the use of the activity stream and consequently the awareness of the activity.

This system will be integrated within the WeSPOT (included in the acknowledgements) system in order to support inquiry learning and it is also open to be tested in other testbeds.

Acknowledgement. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No 318499 - weSPOT project. Katrien Verbert is a post-doctoral fellow of the Research Foundation Flanders (FWO).

References

1. Eales, R.T.J., Hall, T., Bannon, L.J.: The motivation is the message: comparing cscl in different settings. In: Proc. of CSCL 2002, International Society of the Learning Sciences, pp. 310–317 (2002)
2. Santos, J.L., Verbert, K., Govaerts, S., Duval, E.: Addressing learner issues with stepup!: An evaluation. In: Proc. of LAK 2013. ACM (accepted, 2013)
3. Volet, S., Vauras, M.: Interpersonal Regulation of Learning and Motivation: Methodological Advances. Routledge (2013)
4. Williams, K.C., Williams, C.C.: Five key ingredients for improving student motivation. Research in Higher Education Journal 12, 1–23 (2011)
5. Savage, N., Birch, R., Noussi, E.: Motivation of engineering students in higher education. Engineering Education 6(2), 39–46 (2012)
6. Di Serio, Á., Ibáñez, M.B., Kloos, C.D.: Impact of an augmented reality system on studentsmotivation for a visual art course. Computers & Education (2012)
7. Hsiao, I.-H., Brusilovsky, P.: Motivational social visualizations for personalized E-learning. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 153–165. Springer, Heidelberg (2012)
8. Parra, G., Klerkx, J., Duval, E.: What should i read next? awareness of relevant publications through a community of practice. In: Proc. of CHI EA 2013. ACM (accepted, 2013)

9. Wenger, E.: Communities of practice: Learning, meaning, and identity. Cambridge University Press (1999)
10. Govaerts, S., Verbert, K., Duval, E., Pardo, A.: The student activity meter for awareness and self-reflection. In: Proc. of CHI EA 2012, pp. 869–884. ACM (2012)
11. Young, J.R.: Badges earned online pose challenge to traditional college diplomas. Chronicle of Higher Education (2012)
12. Goligoski, E.: Motivating the learner: Mozillas open badges program. Access to Knowledge: A Course Journal 4(1) (2012)
13. Domínguez, A., Saenz-de Navarrete, J., de Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.J.: Gamifying learning experiences: Practical implications and outcomes. Computers & Education 63, 380–392 (2013)
14. Michel, C., Lavoué, E., Pietrac, L.: A dashboard to regulate project-based learning. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 250–263. Springer, Heidelberg (2012)
15. Olson, J.F., Howison, J., Carley, K.M.: Paying attention to each other in visible work communities: Modeling bursty systems of multiple activity streams. In: IEEE SocialCom, pp. 276–281 (2010)
16. Li, I., Dey, A., Forlizzi, J., Höök, K., Medynskiy, Y.: Personal informatics and hci: design, theory, and social implications. In: Proc. of CHI EA 2011, pp. 2417–2420. ACM (2011)
17. Li, I., Dey, A., Forlizzi, J.: A stage-based model of personal informatics systems. In: Proc. of CHI 2010, pp. 557–566. ACM (2010)
18. Verbert, K., Duval, E., Klerkx, J., Govaerts, S., Santos, J.L.: Learning analytics dashboard applications. American Behavioral Scientist (accepted, 2013)
19. Podgorelec, V., Kuhar, S.: Taking advantage of education data: Advanced data analysis and reporting in virtual learning environments. Electronics and Electrical Engineering 114(8), 111–116 (2011)
20. Leony, D., Pardo, A., de la Fuente Valentín, L., de Castro, D.S., Kloos, C.D.: Glass: A learning analytics visualization tool. In: Proc. of LAK 2012, pp. 162–163. ACM (2012)
21. Gutiérrez Rojas, I., Crespo García, R.M., Delgado Kloos, C.: Enhancing orchestration of lab sessions by means of awareness mechanisms. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 113–125. Springer, Heidelberg (2012)
22. Hamari, J., Koivisto, J.: Social motivations to use gamification: an empirical study of gamifying exercise. In: Proc. of ECIS 2013 (2013)
23. Grosseck, G., Holotescu, C.: Can we use twitter for educational activities. In: eLSE08 (2008), http://adlunap.ro/else
24. Zeng, X., Mehdi, Q.H., Gough, N.E.: Shape of the story: Story visualization techniques. In: Proc. of IV 2003, pp. 144–149. IEEE (2003)
25. Dawson, S., Bakharia, A., Heathcote, E.: Snapp: Realising the affordances of real-time sna within networked learning environments. In: Proc. of NLC 2010, pp. 125–133 (2010)
26. Kosba, E., Dimitrova, V., Boyle, R.: Using student and group models to support teachers in web-based distance education. In: Ardissono, L., Brna, P., Mitrović, A. (eds.) UM 2005. LNCS (LNAI), vol. 3538, pp. 124–133. Springer, Heidelberg (2005)
27. Naaman, M., Boase, J., Lai, C.H.: Is it really about me?: message content in social awareness streams. In: Proc. of CSCW 2010, pp. 189–192. ACM (2010)
28. Brooke, J.: Sus-a quick and dirty usability scale. Usability Evaluation in Industry, 189–194 (1996)

Not Interested in ICT? A Case Study to Explore How a Meaningful m-Learning Activity Fosters Engagement among Older Users

Patricia Santos[1], Mara Balestrini[2], Valeria Righi[3], Josep Blat[3], and Davinia Hernández-Leo[3]

[1] Education Department, University of the West of England, Frenchay Campus, Bristol, UK
Patricia.Santosrodriguez@uwe.ac.uk
[2] Intel Collaborative Research Institute on Sustainable Connected Cities, University College London, UK
m.balestrini@cs.ucl.ac.uk
[3] ICT Department, Universitat Pompeu Fabra, C/Roc Boronat 138, 08018 Barcelona, Spain
{valeria.righi,josep.blat,davinia.hernandez}@upf.edu

Abstract. Mobile devices are increasingly being used in lifelong learning. However, while older learners are active members of the lifelong learning system, little research has been aimed at understanding how m-learning can provide them with successful learning experiences. In this paper we address the question if m-learning can foster the acceptance and uptake of mobile technologies among a group of older people unfamiliar with ICT. Following a participatory design approach, 20 participants who were enrolled in a literature course created routes of geolocated questions about a fiction book they were reading and answered them in the real location using the *QuesTInSitu* application. Results indicate that their m-learning acceptance improved as their anxiety around use of technologies diminished. These findings question previous research in which use of mobile technologies tended to increase older users' anxiety and reduced their acceptance of technology. Participants described the experience as playful, enjoyable and useful.

Keywords: M-learning, context-aware, lifelong learning, older learners.

1 Introduction

M-learning (mobile learning) and LBS (Location Based Systems) have mostly been used for supporting learning in formal educational contexts (primary, secondary and high school). In addition, the technologies enabling these systems, such as mobile devices and GPS (Global Positioning System), are increasingly being used in lifelong learning. However, there is a lack of research when it comes to understanding the relationship between older lifelong learners and m-learning. Sharples [1] posits that the use of mobile devices in lifelong learning is *"an extended and holistic process of developing skills and understanding"*, because mobile technology fosters innovative scenarios. Emerging technologies have to be those ones that provide more integration

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 328–342, 2013.

of formal with informal learning, and those ones that support lifelong learning allowing students to access learning at times and places that are personalized and appropriate for them [2-3]. We claim that mobile devices and LBS are suitable tools to support lifelong learning activities also among older users.

Context-aware m-learning activities are aimed at supporting learning in specific places, such as parks, the city, etc. The learner's situation within a context-aware u-learning environment (i) can be sensed, used to conduct an activity, and to offer adaptive supports; (ii) provides supports in the right way, right place, and at the right time; (iii) enables seamless learning within a predefined area; and (iv) is able to adapt the subject content to meet the functions of various mobile devices [4]. Time and location may be the most important and fundamental parameters for recognizing and describing a learner's context [5]. Sharples indicates that situated activities involve learners' reflection, and as a consequence learners have to integrate the current experience with previous knowledge and to construct new interpretations. Accordingly, the technology to support these activities should be mobile, draw on information provided by location-based sensors and related to the learner's situation with the learning task at hand [1].

This paper addresses the relationship between context-aware m-learning and older people (60+), which appears to be an innovative and relevant scenario. Previous research has mostly focused on evaluating the benefits of geolocation technologies and mobile devices to support healthy ageing and independent living [6-8], two key aspects in an ever-growing ageing population. In this paper we aim to understand if m-learning activities can foster the acceptance and uptake of mobile technologies among a group of older people who are not interested in ICT. *How should context-aware m-learning be deployed to provide fruitful outcomes for older learners who are not familiar with ICT?* Within this central research question, we aim to explore 1) if m-learning *in situ* can foster the acceptance of mobile technologies among older people initially not interested in ICT 2) which strategies and combinations of tools can provide an engaging m-learning *in situ* experience.

This project builds on previous research where *QuesTInSitu* (an app to assess learning *in situ*) proved to successfully enable fruitful experiences among young students in formal learning scenarios [15]. We addressed these questions in an experiment conducted with older people recruited from an adult school that offers a variety of free courses supporting lifelong learning, following a co-design approach. During the study, participants worked with the researchers and used mobile devices and LBS with the aim of co-designing an m-learning activity related with their literature course.

The goals were to design and realize an m-learning *in situ* activity that consisted of a route of geolocated questions based on the facts occurred in a novel that they had previously read in a dialogic literary gathering offered by the adult school. Subsequently, the group of learners completed the route *in situ* by answering the questions created by their colleagues, knowing their position, score and feedback *in situ* and in real time using smartphones. The dialogic methodology empowers all the participants to have the same rights in decision making, collaboration and co-design. These methods are in line with the statement of [1] who claims that lifelong learning should be primarily collaborative rather than competitive, and should emphasize the role of the learners as designers of their own learning activities. We propose that adopting a co-design approach to create and perform the context-aware m-learning activity enhanced the m-learning acceptance among our participants and reduced their anxiety towards emerging and unfamiliar technologies. In the following sections, we present the experiment and provide details about the educational activity, the methodology applied. The evaluation and results of the experiment are described. Finally we discuss the main conclusions of the study.

2 Related Work

2.1 Older People and Geo-positioning Systems

Previous works exploring the use of geo-positioning systems by older people have mainly focused on e-health applications. For example, Boulos et al. [8], present a number of projects in which tracking systems are used to support older people with special needs. Many of these systems combine the use of wearable health sensors and GPS embedded in mobile phones to report geographical position and health information to the corresponding professionals.

The use of GPS in leisure, social and learning scenarios has largely been overlooked in HCI research with older people, despite the growing awareness of the importance to promote active, besides healthy, aging [9]. The potential of geo-positioning technologies to support social interactions has been studied by Righi et al. [10] whose initial results revealed a set of potential scenarios for use of geo-positioning technologies, in which sociability, informal and non-utilitarian interactions were proven to be key elements. The use of location-based technologies in leisure settings by older people has also been explored by O'Neil et al. [11], who investigated older people's experience with a location-based mobile multimedia system in a rural nature reserve. Participants reported to be enthusiastic about such technologies but claimed that the system should provide them with richer interactions and contents.

These studies suggest that the use of geo-positioning technologies during an informal m-learning activity is a scenario worth exploring, especially through an approach in which older people take an active role in the design of the activity's contents.

2.2 Technology Acceptance and Older People

Previous research has indicated that technology can be a source of stress among older users. However, if correctly introduced, it can be perceived as useful for improving their physical and mental health [12]. Nycyk & Redsell claim that: *"the effective delivery of computer training to older adults is still a problem. They can feel stressed as their family and peers use technology that they cannot, and they feel bewildered when attempting formal computer training that teaches at a fast pace"*.

An experiment done by Shapira et al. [3] showed that ICT contributes to older adults' well-being and sense of empowerment by affecting their interpersonal interactions, promoting their cognitive functioning and contributing to their experience of control and independence. We have used the study done by Wang et al. [13] as a main reference for our study. These authors highlighted the need to research the factors related to user intention to use m-learning. Their study was based on the unified theory of acceptance and use of technology (UTAUT) [14], which integrates elements across eight models of information technology use. They adapted the UTAUT model and proposed their own research model in order to understand the determinants of m-learning acceptance and to discover if there exist age or gender differences in such acceptance. After conducting an experiment with 330 participants, they concluded that *performance expectancy, effort expectancy, social influence, perceived playfulness, and self-management of learning* were all significant determinants of behavioral intention to use m-learning. In addition, age differences moderate the effects of effort expectancy and social influence on m-learning use intention, and gender differences moderate the effects of social influence, and self-management of learning on m-learning use intention. The determinants factors proposed by Wang et al. in their model have been considered in this study to understand the m-learning acceptance of older people.

2.3 *QuesTInSitu*: An m-learning App for Assessment in Situ

QuesTInSitu is a web-based and mobile app designed and built to support assessment *in situ* activities based on tests [15]. On the one hand, authors can use this system to create their own routes of geolocated questions; on the other hand learners can answer the questions, receive feedback and score *in situ* by using a smartphone and sharing their position. The aim of *QuesTInSitu* is to study the application of handheld devices to open up possibilities for creating more appropriate context-aware activities for lifelong learning. Our view is that new assessment types can be integrated within learning in ways that motivate students and support their learning.

Various experiments in real formal educational contexts (with secondary and university teachers and students) were analyzed to understand the benefits and limitations of using *QuesTInSitu* for doing assessment in situ activities. The results of the scenarios show how *QuesTInSitu* is useful for giving support *in situ* during the realization of an assessment *in situ* activity. The strategy of providing questions, feedback and score in real time and *in situ* is an essential aspect that helps students to better understand the content of the questions of the test-route [15-16].

In this study we use *QuesTInSitu* as an enabler of an innovative *in situ* learning scenario where older learners co-designed, and conducted a context-aware m-learning activity that was meaningful for them.

3 Description of the Experiment

We draw on the study proposed in [13] to understand how *performance* and *effort expectancy*, and *playfulness* affect the participants' m-learning acceptance. We carried out an m-learning activity to engage the older learners (who reported not to be particularly interested in using ICT) in the design of a context-aware m-learning activity by creating a route of questions about a book they had chosen and read.

In the following sections we describe the activity, its design process and the data gathering methods for the evaluation.

3.1 Educational Context

We have conducted the study in an adult educational center[1] that promotes social inclusion of adults through lifelong learning. The center applies an adaptation of the dialogic learning methodology in their courses [17]. This means that all participants' opinions have the same value and are discussed by the group. The participants who took part in the study were all enrolled in a Catalan literature course. There is no teacher or instructor in the course but rather a participant who volunteers to organize the discussion. The book "La plaça del diamant" by Merçè Rodoreda was chosen because its story takes place in a district of Barcelona, the city where the lifelong learning centre is located, and because the participants were very interested in this novel.

3.2 Description of the m-learning Activity, its Design and Implementation Process

After the participants read the book and/or watched the film adaptation, we asked them to come up with questions about specific facts in the story that were related to areas of the district of Gràcia in Barcelona. The questions were defined by the 20 participants in two workshops, which we summarize below. After the co-design of the activity, 11 participants (divided in two groups) participated in the m-learning activity in situ by interacting with smartphones. Finally, in a session with the whole group, we shared a last discussion in order to gather opinions with regards to their personal experience during all the process.

Co-design Workshops

The two workshops (1h 30 min each workshop) held during the same month, aimed at defining the route of geolocated questions related to the selected book. We asked them to voluntarily form two different groups (A, B). During the first workshop the participants in one group had to create questions that would be answered by the participants of the other group during the activity *in situ*. The criterion for selecting the questions was how related they were to a specific physical space: the real position should give the participants a clue to answer the questions. It is important to mention that during this first workshop we did not introduce the use of mobile devices but rather handed paper maps to the participants for them to explore the location of the *in situ* activity and to facilitate the brainstorming of questions.

---

[1] `http://www.edaverneda.org/edaverneda/en`

During the second workshop, participants had to select 10 questions from those proposed in the previous workshop, and create a route by positioning the questions on a specific location on a map. In this case, we provided Apple iPads and Google Maps to re-explore the area of the district selected and find the most suitable locations for the questions. We then uploaded the routes to *QuesTInSitu*.

Activity In-situ

Participants and researchers went to the district of Gràcia in Barcelona. Group A explored and answered the route designed by Group B, and vice versa. Only one mobile phone was assigned to each group because participants refused to carry smartphones individually arguing that they wanted to discuss the questions collaboratively with their team mates. Furthermore, the majority of them reported being afraid of using the devices. Participants had 1h 30 min to plan the best strategy for completing the route, and answering the questions in the correct locations. At the end of the activity a questionnaire was administered and their opinions about the experience were gathered in informal conversations.

Debriefing

A 1h debriefing session was conducted one week after the activity *in situ* took place. In this occasion the data gathered and analyzed were presented and discussed with all the participants of the literature course. Additional comments from participants were also gathered.

Fig. 1 shows the participants during the different phases of the activity (workshop 1 and 2, and the activity *in situ*).

(a) (b) (c)

Fig. 1. (a) Using the paper maps during workshop 1; (b) Interacting with Google Maps and iPads during workshop 2; (c) Activity in situ with smartphones

3.3 Participants

A total of 20 older learners participated in the co-design of the activity and debriefing session, and 11 of them did the activity *in situ* with smartphones. The rest of the participants did not join the activity *in situ* due to mobility/accessibility problems and personal reasons. The average age of the participants was 65 years old. This group was selected as significant because only 1 of the participants had previous experience with mobile technologies; accomplishing the required condition (i.e. older lifelong learners without ICT experience) to understand if the co-design and participation in m-learning activities fostered the participants' mobile acceptance.

3.4 Data Gathering Methods

We adopted a qualitative approach, combining participatory observations and conversations with participants while they were engaging in the activities described above. Three researchers participated in the co-design workshops and debriefing session, and four researchers (two for each group – A, B) participated in the activity *in situ*. During and immediately after the activity the researchers took note of their observations and conversations with participants. Additionally, all the activities were video recorded to further support first-hand observations. A questionnaire was distributed at the end of the activity *in-situ*. The questionnaire, which contained open and closed questions, was aimed to catch positive and negative aspects of the activity and participants' attitudes towards the devices during the activity. However, the questionnaire resulted not to be effective with our participants. In fact, most of them left the open questions unanswered (preferring to comment their experience in face-to-face conversations instead) and answered only a few closed questions. For this reason, in section 4, we do not report extensively on the results of questionnaires. Table 1 summarizes the main data methods adopted in each activity.

Table 1. Data collection techniques

Activity	Type of data	Labels
Co-design workshops	Videos during the workshops (workshop1 and workshop 2). Fieldnotes of 3 researchers' observations and conversations with participants during the activity	[V-ws] [fieldnote]
Activity-in-situ	Videos during the *QuesTInSitu* route. Fieldnotes of 4 researchers' observations and conversations with participants during the activity Questionnaire post activity	[V-gracia] [fieldnote] [Q]
Debriefing	Fieldnotes of 3 researchers' observations and conversations with participants	[fieldnote]

3.5 Data Analysis Methods

Three researchers have analyzed the entire body of fieldnotes and participants' answers to open questions of the questionnaire by using thematic analysis [25]. This consisted in extracting the main themes from the data *corpus*. Themes emerging from the analysis have been discussed by the three researchers until reaching consensus. We identify three themes: i) *Progressive engagement with device*, ii) *Meaningfulness of the activity*, iii) *Collaborative involvement*. Quantitative data extracted from the questionnaires and the videos, have been used to further validate the analysis.

4 Results

In this section we present the partial results collected from the data gathering during the three phases of the experiment. The results are presented according to the three themes emerged from the analysis. The main results of each theme have been codified using a code. These codes are used in the discussion section to support the corresponding proposed implications.

4.1 Progressive Engagement with the Technology - [PET]

We observed that most of the **participants initially refused to interact with the devices** (i.e. mobiles phone in the activity *in-situ*, iPad in the co-design workshop) **suggesting that their *performance and effort expectancy* was negative [PETr1].** This behavior waned as the activities progressed. The following conversation with one participant during the second co-design workshop shows the hesitance of the participant towards using GoogleMap on the iPad to choose the place where to locate the questions: (Participant): *"Do we have to use "it"?(talking about the iPad)."* (Researcher): *As you want, this is not mandatory.* (Participant): *"oh my god! I don't want to use this thing. (Talking to other participant) You have to use it!"* Although that participant never really used the iPad during the activity, a few moments later she asked the researcher to find the street she was looking for and gave instructions about where to move on the digital map. **When the mobile devices were entering in the activity in a natural way, participants were no longer disturbed by their presence [PETr2]**. It is important to note that in the first workshop the activity was explained and participants did not use any technological device to prepare the route of questions. It was during the second workshop when we included the use of iPads. The observations during the second workshop show this change in technology acceptance:*"Participants quickly loose interest in the paper maps and prefer to just talk about the book and the questions interacting with the iPads (comparing this behavior to that in workshop two we can see a difference: iPads seem to be more motivating since users interacted significantly more with them and applied gestures such as swipe, zoom, drag, etc.)"* [V-ws]. Participant's comment supports the previous observation. For example: "*The iPads enabled us to have a clear view of the place, it seemed like we really were at the place!"*[q-user4]. As the following observation indicates participant' acceptance towards the device increased during the activity: "*About halfway through the experience they (participants) relate in a very familiar and comfortable way with the device. They don't seem to be scared of it as they were in the beginning*" [V-gracia]. Participants´ increased confidence (*self-management of learning*) with the mobile device is also supported by the fact that at the beginning of the *in-situ* activity only 3 out of 11 participants offered to carry the smartphone, while after the activity 9 out of 11 participants said that they were willing to carry the device if the activity was repeated. These observations suggest that users **were so engaged in the activity that they started to interact with the technology "naturally"**.

In a similar way, at the beginning of the activity *in situ*, only a few participants accepted to carry the smartphone. Most of them had never used the device before but saw it or heard about it from family members (e.g. one participant commented that her husband had a smartphone and that he used it for many things, including visualizing maps), showing **how *social influence* is an important factor that has to be considered when engaging older learners in m-learning [PETr3]**, see also [17].

4.2 Meaningfulness of the Activities Supported by ICT - [MSI]

The results suggest that their *increased acceptance* of the technology was directly related with their *strong emotional involvement* in the activity *perceiving it as a playful* educational experience [MSIr1]. Engagement was observed during the *in-situ* activity, especially when they discovered that the answers to the questions were correct. Participants displayed different body gestures in those situations: they jumped, they smiled and increased the volume of their voices. In two occasions the participants were so excited to cross the street to answer a question geolocated on the other side of the road that they did not acknowledge that the traffic light was still green for the cars. This observation shows the positive acceptance of the m-learning activity: *"There was a strong emotional involvement during the activity, and especially when they knew that the answers to the questions were correct. Different body gestures were displayed during that moment: they jumped on one leg, they smiled and they increased the volume of their voices. They were so immersed into the activity that once they forgot about the cars traffic and two of them were about to cross the street in a very dangerous way"* [fieldnote]. The fact that the participants were interested in the book used for the learning activity fostered their engagement: during the *in-situ* activity, they chatted about the book and fantasized about the buildings in which the book´s characters might have lived. Also, when participants missed the correct answer of a question they continued discussing it while walking towards the next one. Participants punctuated with an average score of 4,72/6 the usefulness of the activity in situ to better understand the book [q-all_users].

For the participants, the appearance of questions *in situ* was very useful and a confirmation that they were located in the correct position. **The use of prompts (questions and feedback *in situ* and in real time) and location information (e.g. points in a map) engages the participants to observe the physical space, improves their knowledge about the district visited and the association of the real world with the questions augments their point of view about the novel [MSIr2]**. Participants' comments and observations support the previous statements: **"**The *fact that when arriving at the indicated place a question would pop up was very ingenious"* [q-user4]; *"It wouldn't be better to do it on paper. We saw with our own eyes the streets where the characters of the book lived, where the author developed the work."* [q-user1]; *"On the phone we could clearly see those streets that guided you to arrive where you wanted to go"* [q-user4]; "*It was very interesting because the device allowed you to know where you are at all times"* [q-user9]. Participants punctuated with an average score of 5,81/6 the usefulness of the activity in situ to know more about the district visited [q-all_users]. The behavior observed during the activity *in situ* indicates that including paper maps is useful to unite the group and improve the collaboration with the person in charge of the smartphone: "*The printed map was frequently accessed by more than one participant at a time, while the use of mobile devices was more individual*" [fieldnote].

The participants reported that the *context-aware m-learning activity* allowed them to *deepen their understanding of the story*, and have a more realistic view of the scene [MSIr3]. In addition, using *QuesTInSitu* in a real-life context allowed them to improve their knowledge about the visited area, which might have not happened if they had only read the book, *"It wouldn't be better to do it on paper. We saw with our own eyes the streets where the characters of the book lived, where the author developed the work."* [q-user1]. The following partial results confirm the previous statement: participants valued with an average of 5,7/6 the functionality of seeing their positions and questions over the digital map. The use of paper map was punctuated with 5,3/6 points [q-all_users]. These results show that the activity *in situ* improved their learning because they could put in practice their knowledge in a realistic, situated way.

As participants indicated, *QuesTInSitu* enables them to know their current position and find the nearest questions but the **paper-map was a useful complementary instrument to have a global view of the district that they had to explore [MSIr4]**. Observations and comments support the previous results: "One user realizes that the group of dolls on the screen changes according to the geo-positioning of the group: *'The dolls change position according to our movement!'*. She is surprised and excited." [fieldnote]*; "The map (QuesTInSitu map) indicates correctly our situation, it can even be zoomed for a better view"* [q-user1]. We observe how LBS can be combined with traditional paper maps because participants use these instruments to solve different problems during the exploratory task (e.g. the paper map provided a global view of the area, which is difficult to have in the small screen of a mobile phone. The map on the mobile phone facilitated the recognition of their position and allowed zooming functions).

4.3 Collaborative Involvement – [CI]

Regarding the collaborative aspects of the activity, it is worth noting that, as it has been previously indicated, participants worked in groups during the co-design and enactment of the activity. In the design stages, the collaboration between members of the same group is used to (1) understand the use of the technology employed; and (2) discuss the most adequate route and questions in relation to the book. **Collaborative involvement in the co-design of the context-aware m-learning activity encouraged participants' engagement and fostered participants' acceptance of m-learning and LBS [CIr1].** This was due to the fact that all the participants had the same opportunity to share their ideas with their peers, and contribute to the design of the activity. All the participants agreed that doing the activity collaboratively in groups was one of the most positive aspects of the experience. In the *in situ* phase, the group's strategy was to have a mobile-leader and collaboratively find the locations and discuss the questions to answer correctly. The following participant' comments and observations support these statements: *"Among us all we selected the correct answer and shouted of joy when the response was positive"* [q-user2]. The engagement of participants in the co-design process is well expressed by the following comment made by one participant during the debriefing session: *"the activity in situ is not the only thing we have to remember! We worked a lot to create this successful activity! Thinking the questions and locations, designing the routes… we put a lot of effort but also enjoyed a lot"* [fieldnote].

When participants become familiar with the m-learning activity, especially working in teams collaboratively, **their *outcome expectations–performance* and *self-efficacy* improves and their *anxiety* is reduced [CIr2]**, "*Together we found the correct answer, and then we shouted of joy when it was correct"* [q-user2].

In addition, *QuesTInSitu* engaged participants and promoted discussions during the activity: "*When participants miss the correct answer of a question they continue discussing it while walking towards the next one*." [v-gracia]. This behavior shows that **despite the fact that only one person in each group carried the smartphone, all the members of the team were engaged in the activity and *collaborated* in the tasks [CIr3].**

5 Discussion

The aim of this section is to summarize and highlight the implications derived from co-designing a context-aware m-learning activity with older learners to improve their m-learning acceptance. We use the three categories described in the previous section in order to facilitate the organization of implications and main results (using the code references of section 4).

5.1 Progressive Engagement with the Technology

When an emerging technology is used with a group of older learners, it is important to take into account that at the beginning some of them might be scared or uninterested because they have never used it before **[PETr1]**, a similar behavior has been studied with the use of computers and older learners [18]. For this reason, a good start is to ask participants to explain how their social environment (e.g. partners, children, grandsons) use mobile devices **[PETr3].**

As we have observed, a good strategy for designing adequate m-learning activities for older participants is to encourage a collaborative and participative approach throughout the process (conception, development of materials, enactment) **[CIr1, CIr2].** The design of the activity was structured in various stages (see section 2) and the participants had to work collaboratively in groups. The analysis of the results shows that this design helps older learners to understand how and why the technology is employed to perform an m-learning activity, and the organization in phases (workshop 1 and 2) reinforced the involvement of the participants in the creation of the routes and geolocated questions **[PETr2]**. In addition, the involvement of the participants in all the discussions and presentation of results helps them to understand the value of the technology. These data suggest that a traditional, individual experiment, or an evaluation session with m-learning technologies where older people have not been given the opportunity to use them to create learning materials collaboratively, might not be the best approach for providing them with positive and meaningful lifelong-learning experiences.

5.2 Meaningfulness of the Activities Supported by ICT

Previous works showed that usefulness is an important construct to understand the use of ICT by older people. Usefulness has been discussed within the context of mobile phone adoption, video content-generation and social network sites [19-21]. These studies showed that perceived (lack of) usefulness in the technology can encourage (hinder) the uptake of such technology among older people. In our study, although at the beginning the participants did not consider the technology intrinsically useful, they perceived the activity (ICT-supported) as *meaningful*. By the end of the experience, participants showed interest in using the device if a similar scenario was proposed **[MSIr3]**. This result suggests that, although older people might initially be uninterested in the technology **[PETr1]**, using it in an activity which is meaningful for them can contribute to their acceptance creating a strong emotional involvement **[MSIr1]**.

Rosales et al. [22] argued that conducting meaningful activities with older people using unfinished versions of products can contribute to improving the product under development. Our study supports the result of Rosales et al., since it shows that engaging participants in meaningful activity can facilitate the co-design process. At the same time, findings indicate that co-design fostered engagement and motivations among the participants and contributed to add meaningfulness to the activity **[CIr1]**.

On the other hand, the use of LBSs (e.g. *QuesTInSitu*) to perform a context-aware activity with older learners was perceived as beneficial because it allowed them to put in practice the knowledge acquired in the associated real context **[MSIr2, MSIr3]**. Interacting with geographical information helps participants reflect where the questions have to be located (according with the facts of the novel). Participants prefer to combine LBS and paper-maps to perform the *in situ* activity **[MSIr4]**. At the beginning of the in-situ activity, the paper map was often the center of attention and decision-making. However, as the activity unfolded, participants felt more comfortable with the mobile phone and progressively lost interest in the paper map **[MSIr2]**. As the partial results have demonstrated the combination of traditional resources (such as paper maps) with technology increased their confidence.

Based on Wang et al. [13] and Moon & Kim [23] definitions and according to the results described in the above sections, we can claim that older learners perceived the context-aware m-learning activity as playful **[MSIr1]**. These authors defined playfulness as a state of mind that includes three dimensions: the extent to which the individual (1) perceives that his or her attention is focused on the interaction with the m-learning; (2) is curious during the interaction; and (3) finds the interaction intrinsically enjoyable or interesting.

5.3 Collaborative Involvement

The results showed that a collaborative involvement in a m-learning activity (co-design workshops and *in situ* activity) contribute to increase engagement and motivation **[CIr1]**. This result partially confirms previous ethnographical research with older people [24], which showed that social aspects play a central role in their use of ICT. Furthermore, we observed that those individuals (who had some previous experience with ICT, though limited) acted as “ice breakers” for the rest of the group. Observing how a peer used the devices allowed the rest to be introduced to the technology progressively (if there weren’t any “ice breakers”, those who were not willing to carry the mobile phone would have felt forced to use it to start the activity) **[CIr2]**.

6 Conclusion and Future Work

This paper presents a case study exploring how to involve older people who are not intrinsically motivated to learn by using mobile devices, in m-learning activities to foster their acceptance and uptake.

The methodology employed in this study may represent a good practice of m-learning for older users. Our participants thought that their *performance and effort expectancy* improved after doing the activity. Participants found m-learning useful and meaningful because it enabled them to accomplish learning activities putting in practice (in the real world) their knowledge by doing a *playful* activity. As discussed in the results, participants indicated that the use of mobile devices and LBS allowed them to *augment the knowledge* that they previously had about the book. For the design and creation of routes, participants preferred to use iPads to interact with Google Maps because it enabled them to have a *more realistic view* of the learning area. In addition, the QuesTInSitu route mediated by smartphones was a successful activity since participants were immersed finding questions, answering them in the correct place and completing the routes. Their attention was focused on answering the questions correctly, and when one of their answers was incorrect they discussed the other options.

The scenario and findings discussed in this paper contribute to the limited existing body of knowledge around m-learning and older users. On the one hand, it supports that co-design is an adequate approach to enhance m-learning acceptance for older learners. On the other hand, results have shown that participants perceived LBS as useful instruments for learning *in situ*. The meaningfulness of the ICT-based activities in which older people are involved is a key aspect to foster their acceptance of mobile technologies. It is worth noting here that the replication of the same scenario with participants of different profile (e.g. different interests) would not necessarily result in an equally successful experience. In fact, the results suggest that ad-hoc activities should be designed according to the interests of the participants involved in it.

We are currently working on the design and development of a platform that supports the learning and sharing of knowledge through games created by older people [22]. The platform embraces the concept of *meaningful activities* by allowing users to create their own game. In the future we plan to further explore m-learning with older people through the use of the aforementioned platform to gain insights on which other scenarios can be used to improve the experience of older participants during a m-learning *in situ* activity. With further results, it will be possible to propose good practices for lifelong learning applying context-aware m-learning for older users.

Acknowledgments. This work has been partially sponsored by the Life 2.0: Geographical positioning services to support independent living and social interaction of older people (CIP ICT PSP-2009-4-270965), the Spanish Ministry of Economy and Competitiveness EEE Project (TIN2011-28308-C03-03) and WorthPlay (funded by Fundación General CSIC and la Caixa Obra Social). The authors would also like to thank the GTI group and the Escola d'Adults de la Verneda-Sant Martí (Àgora).

References

1. Sharples, M.: The design of personal mobile technologies for lifelong learning. Computers & Education 34(3-4), 177–193 (2000)
2. Johnson, L., Smith, R., Willis, H., Levine, A., Haywood, K.: The horizon report: 2011. The New Media Consortium, Austin (2011)
3. Shapira, N., Barak, A., Gal, I.: Promoting older adults' well-being through Internet training and use. Aging & Mental Health 11, 477–484 (2011)
4. Hwang, G., Tsai, C., Yang, S.J.H.: Criteria, strategies and research issues of context-aware ubiquitous learning. Educational Technology & Society 11(2), 81–91 (2008)
5. Lonsdale, P., Baber, C., Sharples, M., Arvanitis, T.N.: A context awareness architecture for facilitating mobile learning. Learning with Mobile Devices: Research and Development, 79–85 (2004)
6. Calvo-Palomino, R., de las Heras-Quirós, P., Santos-Cadenas, J.A., Román-López, R., Izquierdo-Cortázar, D.: Outdoors monitoring of elderly people assisted by compass, GPS and mobile social network. In: Omatu, S., Rocha, M.P., Bravo, J., Fernández, F., Corchado, E., Bustillo, A., Corchado, J.M. (eds.) IWANN 2009, Part II. LNCS, vol. 5518, pp. 808–811. Springer, Heidelberg (2009)
7. Helal, S., Winkler, B., Lee, C., Kaddourah, Y., Ran, L., Giraldo, C., et al.: Enabling location-aware pervasive computing applications for the olderly. In: Proceedings of the First Pervasive Computing Conference, Texas, USA, pp. 351–358 (2003)
8. Boulos, M., Anastasiou, A., Bekiaris, E., Panou, M.: Geoenabled technologies for independent living: Examples from four european projects. Technology and Disability 23(1), 7–17 (2011)
9. World Health Organization Active Ageing: A Policy Framework. Geneva (2002)
10. Righi, V., Malón, G., Ferreira, S., Sayago, S., Blat, J.: Preliminary Findings of an Ethnographical Research on Designing Accessible Geolocated Services with Older People. In: Stephanidis, C. (ed.) Universal Access in HCI, Part II, HCII 2011. LNCS, vol. 6766, pp. 205–213. Springer, Heidelberg (2011)
11. O'Neill, E., Collomosse, J., Jay, T., Yousef, K., Rieser, M., Jones, S.: Older User Experience. Vehicular Technology Magazine 5(1), 31–38 (2010)
12. Nycyk, M., Redsell, M.: The role of computer tuition in community health: A grounded theory approach. Aging International 31, 296–308 (2006)

13. Wang, Y., Wu, M., Wang, H.: Investigating the determinants and age and gender differences in the acceptance of mobile learning. British Journal of Educational Technology 40(1), 92–118 (2009)
14. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS Quarterly 27(3), 425–478 (2003)
15. Santos, P., Pérez-Sanagustín, M., Hernández-Leo, D., Blat, J.: *QuesTInSitu*: From tests to routes for assessment *in situ* activities. Computers & Education 57(4), 2517–2534 (2011)
16. Pérez-Sanagustín, M., Santos, P., Hernández-Leo, D., Blat, J.: 4SPPIces: A case study of factors in a scripted collaborative-learning blended course across spatial locations. International Journal of Computer-Supported Collaborative Learning 7(3), 443–465 (2012)
17. Sánchez-Aroca, M.: La Verneda-Sant Martí: A school where people dare to dream. Harvard Educational Review 69(3), 320–335 (1999)
18. Karavidas, M., Lim, N.K., Katsikas, S.L.: The effects of computers on older adult users. Computers in Human Behavior 21(5), 697–711 (2005)
19. Conci, M., Pianesi, F., Zancanaro, M.: Useful, social and enjoyable: Mobile phone adoption by older people. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 63–76. Springer, Heidelberg (2009)
20. Sayago, S., Forbes, P., Blat, J.: Older people's social sharing practices in YouTube through an ethnographical lens. In: Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers, pp. 185–194. British Computer Society (2012)
21. Righi, V., Sayago, S., Blat, J.: Older people's use of Social Network Sites while participating in local online communities from an ethnographical perspective. In: Proceedings of CIRN 2012 Community Informatics Conference: 'Ideals meet Reality', Prato (Italy), November7–9 (2012)
22. Rosales, A., Righi, V., Sayago, S., Blat, J.: Ethnographic techniques with older people at intermediate stages of product development. In: Proc. How to Design Touch Interfaces for and with Older Adults: Identification of Challenges and Opportunities, Nordic CHI (2012)
23. Moon, J., Kim, Y.: Extending the TAM for a world-wide-web context. Information & Management 38(4), 217–230 (2001)
24. Sayago, S., Blat, J.: An ethnographical study of the accessibility barriers in the everyday interactions of older people with the web. Universal Access in the Information Society 10(4), 359–371 (2011)
25. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative Research in Psychology 3, 77–101 (2006)

Exploring LogiAssist – The Mobile Learning and Assistance Platform for Truck Drivers

Maren Scheffel[1], Uwe Kirschenmann[1], Andreas Taske[1], Katja Adloff[2], Maik Kiesel[3], Roland Klemke[3], and Martin Wolpers[1]

[1] Fraunhofer Institute for Applied Information Technology FIT, Schloss Birlinghoven, 53754 Sankt Augustin, Germany
{maren.scheffel,uwe.kirschenmann}@fit.fraunhofer.de, {andreas.taske,martin.wolpers}@fit.fraunhofer.de

[2] Julius-Maximilians-Universität Würzburg, Institut für Psychologie Röntgenring 10, 97070 Würzburg, Germany
katja.adloff@stud-mail.uni-wuerzburg.de

[3] Humance AG, Goebenstraße 10-12, 50672 Köln
{maik.kiesel,rklemke}@humance.de

Abstract. Within the last few years mobile learning has become very popular. Many projects and case studies explore its usefulness in the educational sector, i.e. at schools and universities. Few, however, deal with work place learning. LogiAssist is an educational assistance and learning platform for truck drivers. This paper first explores how eager truck drivers are to engage in online learning and the usage of an assistance system. This is juxtaposed with first results of the actual usage of LogiAssist. In a second step, we explore how the system was used and how the usage can be interpreted from a learning analytics point of view.

Keywords: mobile learning, contextual learning, self-regulated learning, assistance systems, learning analytics.

1 Introduction

With the widespread use of mobile devices, mobile learning has become a very powerful approach for learning and teaching with a steadily growing market. Nowadays, schools, universities, healthcare and other organisational systems already benefit from mobile learning in a large number of scenarios (cf. [1–4]). Mobile learning is especially interesting because the new devices used have the potential to change the way users behave, the way users interact with each other and their attitude towards learning [5]. The key to enable mobile learning is its contextualisation: (1) the learning material is integrated in everyday life situations; (2) clearly defined goals and objectives that reflect daily practice and usage scenarios help learners to become and stay motivated; and (3) learning bases on the interaction of learners [6]. Furthermore, mobile learning must not only support a sound pedagogical concept but must also aim to satisfy the needs of the respective target group [7].

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 343–356, 2013.

This paper focuses on truck drivers as the target group for mobile learning. The presented insights have been obtained by applying LogiAssist in such scenarios. LogiAssist[1] is a mobile assistance and learning platform for truck drivers developed by a German consortium and funded by the NRW Bank in the Ziel 2 sponsorship program. Based on the aforementioned principles for mobile learning, LogiAssist is designed to assist German truck drivers in their daily work while also providing suitable learning activities. Additionally, it helps them to prepare for more demanding and prospective tasks, thereby supporting further qualification. In order to meet the needs of the druck driver community, the acceptance and usage of an elaborated prototype have been evaluated with German truck drivers from different trucking companies. Introducing a learning system like LogiAssist requires a slightly different approach than when dealing with university students as truck drivers are on average not a computer affine group. LogiAssist is still under development, thus the first results provide indications on how to improve the system further. Several aspects will be explored in this paper: first, are truck drivers open-minded towards a comprehensive learning and assistance platform? Second, can an assistance-oriented application be helpful and can it support truck drivers even though they do not make heavy use of a dedicated learning platform? Third, can learning analytics methods be useful for getting an overview of users' behaviour within the application? As the usage of LogiAssist is still relatively sparse, question two cannot be explored to its full degree yet.

The rest of the paper is structured as follows: First, we will describe the background situation for truck drivers and related work (section 2) followed by an explanation of the architecture of LogiAssist (section 3). After that we provide the results of the questionnaire in a compact way (section 4) and then take a closer look at the actual usage of the system (section 5). The usage data is provided in descriptive statistical terms. In addition several analyses are made that give some insights to the usage data in terms of learning analytics. Finally, we will conclude (section 6) by discussing the data and analysing the actual development stage of LogiAssist.

2 Background

Truck drivers are a very interesting target group: Not only are they very mobile but they have to meet rising demands like tighter timetables, ever higher capacity utilisation and traffic that is continuously increasing (70%-80% from now until 2025 [8, 9]). Additionally, demographic changes force them to work longer, i.e. more hours as well as more years. In general, truck drivers are thus under a considerable strain with high stress levels. Over and above, technical improvements such as telemetric systems, digital speedometers, assistance systems and changing regulations for load securing and frontier crossing force the drivers to learn and integrate new behaviours into their daily routines [8].

[1] `http://www.logiassist.de/`

In Germany, the law on the qualification of commercial drivers (BKrQG, Berufskraftfahrer-Qualifikations-Gesetz) regulates transportation procedures as well as the qualification of truck drivers. According to the BKrQG exams have to be taken every five years. The aim of these exams is to improve the safety, to facilitate economical driving styles and to keep a common qualification standard for drivers in the EU. For many truck drivers this is an obstacle that is not overcome easily. Especially newcomers with low school levels or language difficulties due to a foreign background often fail to pass BKrQG exams.

Few projects already explore online learning solutions for truck drivers, or more generally for hands-on jobs with a high need for mobility. One of these is called BLOOM (Bite-sized Learning Opportunities On Mobiles)[2]. It aims to deliver work-place training for the logistics and passenger transport sectors using a variety of mobile phone technologies. The project conducted case studies examining best practices for learning, when to learn and what material taxi drivers found interesting. The project concluded that blended learning approaches are an important aspect to bring taxi drivers together. Another closely related project to LogiAssist is "Mobile Learning" by the Fernuniversität Hagen [10]. The aim of the project was to explore new ways of mobile learning. Therefore, truck drivers were equipped with netbooks to access their mobile learning content. The learning topics focussed on health and safety issues. Drivers were free to learn whenever they wanted . Results showed that drivers mainly learn in their truck when they are alone and during breaks [11].

LogiAssist differs from both projects in that it integrates daily needed functions and operates as an assistance system rather than a learning system. The learning component is integrated as one element of many to ensure that the truck drivers use the system on a daily basis and thereby get used to it (cf. results from the "Mobile Learning" project). Hence, learning using the LogiAssist system becomes just one of the various tasks for which they use the system. Truck drivers are a specific user group: they are usually alone, have to handle very concrete situations and are interested in quick feedback [12]. Consequently, the entry level of LogiAssist to truck drivers is significantly lower than a system provided with the sole purpose of learning. In line with this approach, LogiAssist learning support is facilitated through intuitive and simple to use tools. The pedagogical approach is formulated through the learning resources (provided by publishing houses) and a simple, usage-based recommender system.

3 The LogiAssist System

LogiAssist was developed as a shippable product that addresses and offers solutions for typical daily problems of truck drivers. It can be used by small as well as larger companies and can be fed with contextual and important information that is relevant to the drivers. Its generic framework integrates social network features to exchange news and communicate, rate and comment on events and other activities. Drivers can get navigational hints or use checklists to perform

[2] http://www.bloom-eten.org/

their routine tasks. Learning is integrated in assistance functions as well as in dedicated learning sections. As such the approach is very broad and provides the basic architecture for a sound and comprehensive assistance and learning platform. It has been developed as an application for smartphones and thus can even be valuable for individual truck drivers whose company does not equip them with company devices. As truck drivers need situational relevant and quick feedback and because learning material such as lengthy tutorials might be more intimidating than they are helpful, assistance and community design aspects are both integrated in LogiAssist to support setting up informal learning experiences.

In detail, LogiAssist consists of nine modules (see Figure 1): The *Map* enables the users to see their own location and displays points of interest, e.g. fuel stations, police stations, pharmacies that are close by. *POI/Events* handles the creation and maintenance of POIs. Here, truck drivers can enter interesting POIs as well as recommendations and ratings or write reports about traffic jams or building sites. *DocStop* provides truck drivers with information about medical drop-in centres in case of emergencies. This module can help to find a doctor close to the current location and provides information about office hours, emergency numbers and languages spoken at the respecting facilities. In case of an emergency a button can be clicked transmitting a short message to a specified contact. The module *Ausbildung* (i.e. Education) holds educational information and a number of technical and social topics as audio, text, and video files. Two important factors for learning success are exchange and collaboration [6]. Therefore, the platform *Community* serves as the message exchange centre and provides an overview of connected friends, an inbox folder and a calendar for important dates. Quick and direct help is given by the *Gefahrgut* application which is an easy to use assistance system for dangerous goods where users can easily look-up UN-Numbers for their load. *Vokabeln* offers users vocabulary learning lists for different languages. The last two modules are *Formulare*, an area for forms, checklists and questionnaires, and *Profil*, enabling the user to create their own profile page and to update password and photo.

Except for *DocStop*, *Vokabeln* and *Gefahrgut* all modules were developed using PhoneGap with HTML and JavaScript in order to make LogiAssist a cross-

Fig. 1. LogiAssist Module Overview

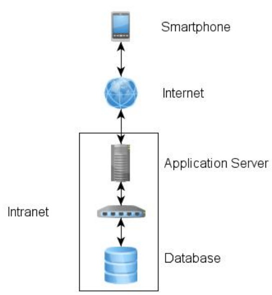


Fig. 2. LogiAssist Tier Architecture

platform system. For those six modules a logging framework was developed that tracks user actions within the applications. *DocStop*, *Vokabeln* and *Gefahrgut* are separate applications that have been integrated into LogiAssist but no usage data is available for them. The architecture of LogiAssist consists roughly of three tiers (see Figure 2): via the mobile device the user connects to the application server in a protected area. From there all necessary and context-based data is downloaded to the user after a successful login.

4 Attitude towards Mobile Learning

In order to measure the truck drivers' attitude towards technology and mobile learning, the logistics company Werner Ruploh KG[3] did a questionnaire with seven of its drivers. In addition, a second questionnaire was made available to truck drivers as a part of the LogiAssist system and could be answered directly within the application. This second questionnaire was answered by 88 participants between February 2012 and January 2013. The aim of the questionnaires was to extract information about the attitude of the drivers towards technology, learning and assistance systems. The detailed evaluation results can be downloaded from the LogiAssist website[4].

Table 1. Most important results from the LogiAssist questionnaire

Item	**Highest Rated Option**	**Results in %**
Mainly used electronic device on the way?	Smartphone	95.4
Which information would be valuable for your everyday job?	Traffic jams and construction sites	92.0
Would you prepare for educational courses when on the way?	Yes	93.2
What is for you the easiest way to learn?	Practical examples	83.0
How important are tests to measure learning success?	Important	90.0
Which topics are most important for you?	Regulations for driving and breaks	75.0

70% of the participants of the LogiAssist questionnaire were between 18 and 40 years old, and only 30% were above 40 years of age. 41.5% had driving experience up to 5 years and 22% up to 10 years, with rates decreasing for more experience. The load capacity of the vehicles was dominated by 25t-40t (40%) followed by the weight class above 40t (15%). Average driving distance a week was 2185

[3] http://www.ruploh.de/

[4] http://www.logiassist.de/web/guest/forschung

kilometres. As distances vary considerably, the standard deviation (SD) was computed at 1178km. The most important results are listed in table 1 and show the highest rated options. It is obvious that the drivers show the willingness to use a learning and assistance system. From a marketing point of view it is interesting that drivers are willing to pay for useful help: they would pay an average of 9.42€a month for a useful application that supports them in learning and assistance (SD=10.51€, maximum(max)=50€).

The evaluation done by the Werner Ruploh KG included seven drivers and was conducted in the beginning of 2010 in the form of face to face interviews. In contrast to the LogiAssist questionnaire, 85% of the participants were above 40 years of age. 57% had driving experience of more than 31 years, and only one was a beginner. The load capacity of the vehicles was dominated by 25t-40t (40%) followed by the weight class above 40t (15%). Average driving distance per week was 2100 kilometres (SD=848 km). Again, the most important results are given in table 2. On average the drivers stated to be willing to pay 62.5€a month for a useful learning and assistance application with a maximum of 175€(SD=73.48€). In addition 71.4% answered that their only qualification to work as a truck driver was the acquisition of the driving license.

The questionnaire results show that truck drivers are open minded towards learning and assistance systems. This is true for the rather older subpopulation of the Werner Ruploh KG as well as the younger participants that downloaded LogiAssist on their own. Even though the Ruploh sample is relatively small, results can be termed representative as they replicate the findings of the Fernuniversität Hagen [11, 13]. From a financial point of view a mobile solution for truck drivers seems to be a worthwhile investment. As questionnaire statements and actual behaviour can differ, we also analysed the usage data collected with the LogiAssist system.

Table 2. Most important results from the Werner Ruploh KG questionnaire

Item	**Highest Rated Option**	**Results in %**
Do you have experience with smartphones?	Yes	28.6
Which electronic device are you experienced with?	PC	71.4
Which information would be valuable for your everyday job?	Traffic jams and construction sites	92.3
Would you prepare for educational courses when on the way?	Yes	85.6
What is for you the easiest way to learn?	Practical and educational examples	85.6
How important are tests to measure learning success?	Important	85.6
Which topics are most important for you?	Economic driving	100.0

5 Usage Data Analysis

The positive questionnaire results were compared to those of the usage data analysis of the LogiAssist application to disclose (a) the acceptance of the application and (b) its actual usage. The evaluation period lasted from June 2012 until March 2013 and includes all users that downloaded and tested LogiAssist.

5.1 Descriptive Statistics

LogiAssist has 245 registered active users. Most of the active users ran LogiAssist only a few times. Figure 3 depicts how many sessions the users had. Over 80% only had one to five sessions with exponentially decreasing percentages for those with more than one session. The overall average number of clicks per user was 54. Including all the sub menus there are 36 pages available in LogiAssist. An average user therefore made 1.5 complete run-throughs of the application.

The average time spent on the application per month was 4.17 minutes. Over time, this did not change much (SD=0.87 min). The average session duration was 4 minutes (SD=5.2 min) and the average time spent on a page was 14 seconds. In addition the distribution of the average time spent according to the first level hierarchy categories is listed in Figure 4. These differences show that some categories were merely browsed through, while others were regarded more intensely.

When these results are compared to the group of the Top 5 power users with highest session counts, the results are not much different: The average number of clicks per power user was much higher (270) but the average amount of time spent per page was also 13.6 seconds which is closely the average of the time spent for each category (14.3 sec). Although power users spent more time with the application, their browsing behaviour was not different from the overall average.

The descriptive analysis of the collected usage data shows that LogiAssist was tested and used only superficially so far. Although many users played with it, very few undertook the effort to examine it more closely. This contrasts with the very positive questionnaire results reported in section 4 but was to be expected. Three main factors are responsible for this. Firstly, LogiAssist does provide a powerful architecture but individualised data like checklists, timetables or routing plans are not integrated yet. Secondly, as companies provide drivers with this important information, tests and important hints for their routes, the application will become more helpful. Also, LogiAssist will be more valuable the more users rate, comment and enter information. This is a typical cold start problem. Thirdly, LogiAssist is still under development and some parts offered in the menu had thus not been implemented yet or were difficult to use. The process analysis offers hints at what users failed to see and how LogiAssist could be improved and reveals the main workflows of the application. To this end two means of analysis were used: usage graphs and key action extraction. Although these analyses were done with the reduced data provided (i.e. not from all modules), their usefulness for learning analytics shall be documented. Application developers can use this information to see, where things went wrong. Further

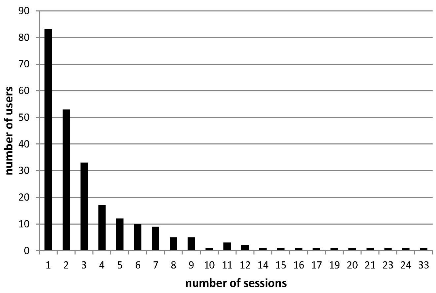


Fig. 3. Num. of sessions per num. of users

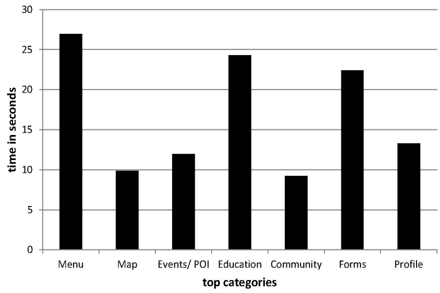


Fig. 4. Average time per top category

qualitative feedback from users is shortly evaluated in the conclusion. This shall give more insight, in why the application was not fully used.

5.2 Usage Graph Analysis

As described in section 4 the actual usage of LogiAssist is sparse. Nevertheless, we provide some methods to gain deeper insight in how LogiAssist was used and could be used. The usage graph analysis reveals which steps users followed to navigate within the application. Usage graphs can help to uncover where working flows stopped and users had to return to previous actions. By design, usage graphs help in getting a very quick overview of the mainly used workflows within an application. As such they are helpful in analyses with more data as well.

In the graphs for this sections nodes indicate the application pages and the connections between the nodes (edges) show how often users moved from one page to another. As the sub-hierarchy is relatively complex, especially in the category of *POIs/Events* and *Community*, this visualisation can render new insights. In the following sub-sections graphs are displayed that show the direction of navigation and the frequency in which these steps were taken. The frequency of transitions is displayed as thickness of the connecting edges and the arrow size. The summed up clicks for pages are indicated by the diameter of the circles. These graphs sum up the activity of all users or the amount of users that browsed into the sub-hierarchy. For convenience it is said that a decrease in clicks in sub-hierarchies mirrors a decrease in user activation. In most case this should be true.

Education. The education area (i.e. *Ausbildung*) has a very clear usability flow: From the Menu page (mainMenuPage) the user gets to the overview of learning units (learningUnitListPage). From there the navigation leads to the several learning objects (learningObjectListPage) as video or text files (textLoPage, videoLoPage). The usage graph of the education category in Figure 5 shows that users did get through to the list of learning objects that belongs to a learning unit. There, however, 35% returned without clicking on the actual learning content. When they did look at learning content, it was mostly text files. Videos

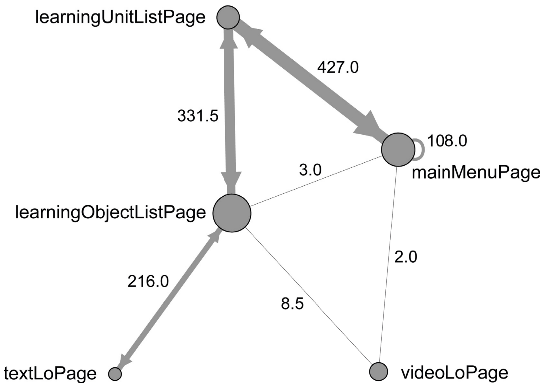


Fig. 5. Usage graph for the module *Ausbildung*

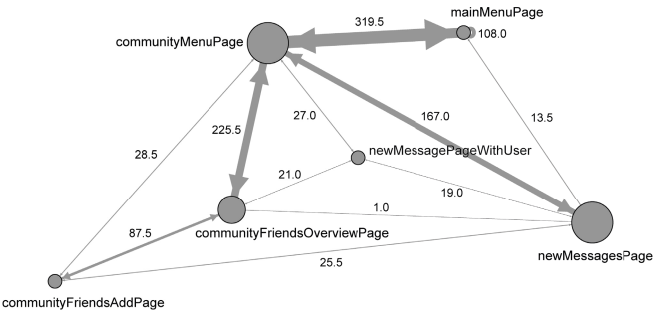


Fig. 6. Usage graph for the module *Community*

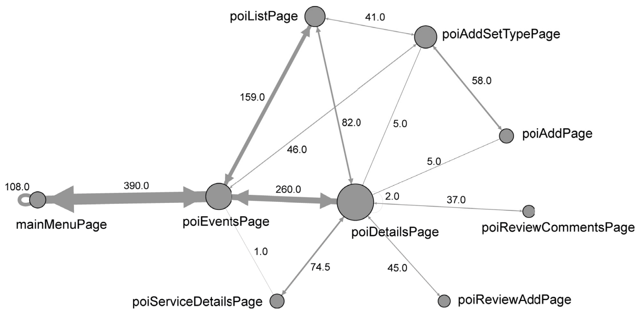


Fig. 7. Usage graph for the module *POIs*

were just looked at by 2% of the users. This could be due to the fact that videos must be downloaded separately and that drivers are more interested in quick feedback rather than in lengthy learning material. The average time spent for texts was equal to that for the average of pages: 14 seconds. Users that watched videos, spent some time on it, but usually did not watch the whole video: average playing time for videos was 38 seconds.

Community. The work flow in the community area is more complex. Users get from the main menu to the *Community* page. There they have the options of checking their calendar (functionality not yet included), getting an overview of their contacts within LogiAssist and writing new message from their message box folder. 52% of the users that browsed to the *Community* module did check out the inbox folder (newMessagesPage). 70% looked at the friends page (communityFriendsOverviewPage), but of them only 38% really tried to add friends (communityFriendsAddPage). Direct ways to add friends or write a message from the community menu (newMessagePageWithUser) were chosen by only 9% and 8% of the users respectively.

POIs. The third work flow we present here is the most complex one. In the *POI* module users have the possibility to comment on POIs, add new POIs, rate POIs etc. Users that browse to the *POIs* module get an overview of all events and POIs that are entered into the system. By clicking an interesting POI, the user gets detailed information (poiDetailsPage). 66% of the users did look up details to a POI. From this subgroup 7% and 17% respectively commented on or rated a POI. 2% added a POI from the details page (poiAddPage). The percentage was higher for those users that added a page from the POI template page (poiAddSetTypePage, 70%).

5.3 Key Action Extraction

Key action extraction is a concept borrowed from text analysis: instead of extracting key words from a text, we extract key actions from usage data. Key words significantly describe the content of a text and thus form a useful semantic representation of it [14], i.e. key words capture the essence of a text and deliver essential information about it. Along those lines, we take key actions to significantly describe the content of a session of user actions and form a useful semantic representation of it. Key actions are indicative of what a user has been doing and present essential activities. Because of this, they complement the usage graph analysis where users are only indirectly addressed.

The n-gram approach [15] we use for the analysis of the usage data collected by LogiAssist has already successfully been applied to usage data collected from students in programming courses and its results were used by teachers to evaluate their courses and gain insights in to the students behaviour [16][5]. Key action extraction can thus help to distil data and to identify reoccurring patterns of action sequences by one or more users.

[5] A detailed description of the key action extraction n-gram algorithm can be found in [17].

In the following we show three categories of key action sequences extracted from the LogiAssist usage data: simple 3-step sequences done by many different users, a little more advanced sequences of 4 to 5 steps done by several different users and finally a very complex sequence done repeatedly by one user. It is also demonstrated that key action extraction is very helpful in identifying pitfalls in the application.

Simple Sequences (3 steps). Figure 8 shows typical browsing behaviour of users. From the main menu they jump to a certain topic and back to the main menu. The sequence gives a representative picture of the typical usage of LogiAssist.

Medium Sequences (4-5 steps). These behaviours were shown by at least four to five different users. In Figures 9 and 10 we have some slightly more sophisticated user tasks: In the first case a user checks out his friend overview page with his enlisted friends. Then he adds a friend from the registered users and goes back to the overview page and the community menu page. A similar case is given in Figure 10, where a user looks at the *POIs/Events* overview page. He then takes a look at the POIs only, goes back to the overview page and from there to the main menu. Figure 11 reveals a typical problem with the application that occurred repeatedly. As LogiAssist requires a login, a fast click back from the main menu forces the user to re-enter with username and password. From there all needed local information is reloaded from the server which takes 10-15 seconds. This is a break in the information flow and could lead to earlier drop-out.

Complex Sequences (10 steps and more). A sophisticated usage chain done by one user is shown in Figure 12. This chain happened more than six times and displays a certain feedback cycle. The user was on the *POI/Events* overview page and from there took a look at the *Map* to see the closest POIs. Then he returned to the *POI/Events* module (poiAddSetTypePage) to add a POI (poiAddPage). When this was done, the POIDetailsPage was checked to see if the POI is listed. Albeit this sequence was interrupted by a quick click on the poiAddSetTypePage that displays all templates that can be used for adding a POIs, because in the next sequence the user returned to the *POI/Events* overview page, checked the *Map* again, presumably to see the previously entered POI at hand and the cycle started over again.

Usage graphs and key action extraction are helpful in identifying user behaviour and transitions within the application. The analyses confirm the results of the descriptive statistics in section 5.1. Very few users made the effort to examine LogiAssist in more depth. Especially the usage graphs show the complex arrangement of pages and their possible navigational connections. The key action extraction displayed reoccurring workflows and some pitfalls users encountered regularly. Complex activities show, how LogiAssist can be used if tested to its full potential.

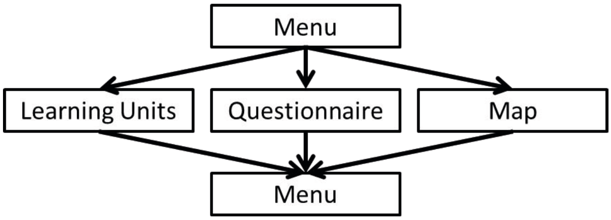


Fig. 8. Simple sequence

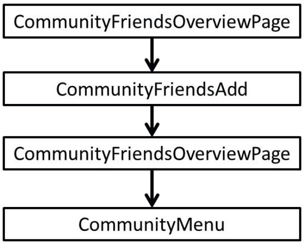


Fig. 9. Medium sequence a

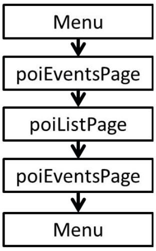


Fig. 10. Medium sequence b

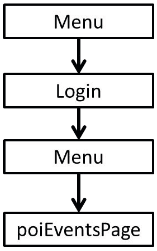


Fig. 11. Medium sequence c

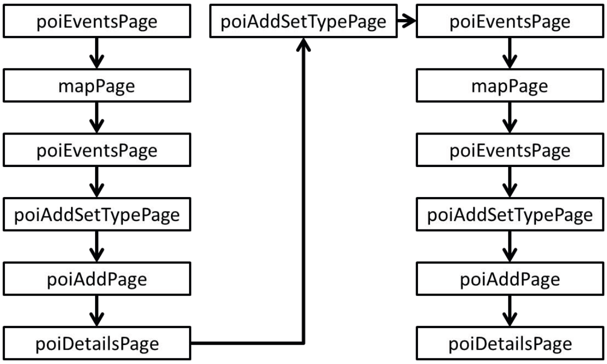


Fig. 12. Complex sequence

6 Conclusion

LogiAssist is a powerful assistance and learning application for truck drivers. Its architecture and design enable a comprehensive support in everyday life. We also found that truck drivers want to be supported with electronic devices and show interest in learning and are even willing to pay for the support. Interestingly, the acceptance was not age-dependent - both age groups above and below 40 years of age stated that combined assistance and learning systems like LogiAssist would be highly interesting to them. More importantly, they required a tight integration into their daily work life. Examples could be the automatic provision of driving assignments by the trucking company, loading instruction support in relation to their educational level, etc.

However, in its actual stage few users made use of the whole potential that LogiAssist offers as was shown by the descriptive analyses (section 5.1). The slow uptake bases on a not user-friendly interface design and clumsy user guidance. The problem of quickly falling back to the login page and having to reload local data is irritating. In respect to the *Education* module, truck drivers expect small learning resources rather than extensive courses. Furthermore, learning material must address the current specific situation rather than general themes. The results from sections 5.2 and 5.3 show the potential of LogiAssist but also its pitfalls.

The above mentioned problems have led to the redesign of LogiAssist. Version 3.0 of LogiAssist comes up with a new design and better performance values. Already now, user reactions on the Google Play Store show a high interest in the application and its assistance plus learning approach[6]. Until mid June 2013 the LogiAssist app has been downloaded 1525 times with an average rating of 3.5 out of 5. Other parts like the *Gefahrgut* app are highly acclaimed[7]. The app had 13357 downloads until mid June 2013 with an average rating of 4.6 out of 5. It seems, therefore, that the experience from the prototype is well taken into account. Further publications will report additional findings when drilling down on specific modules of LogiAssist, e.g. the *Education* module where available learning resources are currently added.

Acknowledgements. The research leading to these results has received funding from the European Union: Investition in unsere Zukunft – Europäischer Fonds für regionale Entwicklung.

References

1. Kearney, M., Schuck, S., Burden, K., Aubusson, P.: Viewing mobile learning from a pedagogical perspective. Research in Learning Technology 20 (2012)
2. Kinash, S., Brand, J., Trishta, M.: Challenging mobile learning discourse through research: Student perceptions of Blackboard Mobile Learn and iPads. Journal of Educational Technology 28(4), 639–655 (2012)
3. Kukulska-Hulme, A., Sharples, M., Milrad, M., Arnedillo-Sanchez, I., Vavoula, G.: The genesis and development of mobile learning in Europe. In: Parsons, D. (ed.) Combining E-Learning and M-Learning: New Applications of Blended Educational Resources, pp. 151–177. Information Science Reference (an imprint of IGI Global), Hershey (2011)
4. Bartelsen, J.: Mobile Learning: Ein Überblick, Arbeitspapiere der Nordakademie, 2011-06 (2011)

[6] https://play.google.com/store/apps/details?id=de.humance.android.logiassist.phon egap

[7] https://play.google.com/store/apps/details?id=de.humance.android.gefahrgut

5. Traxler, J.: Defining, Discussing, and Evaluating Mobile Learning: The moving finger writes and having writ.... International Review of Research in Open and Distance Learning 8(2) (2007)
6. Parsons, D., Ryu, H., Cranshaw, M.: A Study of Design Requirements for Mobile Learning Environments. In: Proceedings of the IEEE International Conference on Advanced Learning Technologies 2006 (ICALT), pp. 96–100 (2006)
7. Schmitz, B., Czauderna, A., Klemke, R., Specht, M.: Game Based Learning for Computer Science Education. In: Proceedings of Computer Science Education Research Conference, pp. 81–88. ACM, Heerlen (2011)
8. Bergrath, J.: Rechnung ohne Fahrer. Fernfahrer 5, 6–8 (2011)
9. Helmrich, R., Zika, G. (eds.): Beruf und Qualifikation. BiBB-IAB-Modellrechnungen zu den Entwicklungen in Berufsfeldern und Qualifikationen bis 2025. Bonn (2010)
10. Özcelik, F., Trojahner, I.: Mobile Learning für Berufskraftfahrer im Fernverkehr. Waxmann Verlag (2011)
11. Karpfen, B.: Mobile learning für Berufskraftfahrer, `http://www.slideshare.net/benjaminkarpfen/mobile-learning-fr-berufskraftfahrer-fernuni-hagen`
12. Roetting, M., Huang, Y.H., McDevitt, J.R., Melton, D.: When technology tells you how you drive - Truck drivers' attitudes towards feedback by technology. Transportation Research Part F: Traffic Psychology and Behaviour 6(4), 275–287 (2003)
13. Dietrichs, J.: Mobile Learning an deutschen Universitäten: Eine empirische Analyse von Einstellung und Akzeptanz zu mobilem Lernen bei Studierenden, Grin, München (2011)
14. Rose, S., Engel, D., Cramer, N., Cowley, W.: Automatic Keyword Extraction from Individual Documents. In: Berry, M.W., Kogan, J. (eds.) Text Mining, pp. 1–20. John Wiley & Sons Ltd. (2010)
15. Hulth, A.: Improved Automatic Keyword Extraction Given More Linguistic Knowledge. In: Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, EMNLP 2003, pp. 216–223. Association for Computational Linguistics, Stroudsburg (2003)
16. Scheffel, M., Niemann, K., Leony, D., Pardo, A., Schmitz, H.-C., Wolpers, M., Delgado Kloos, C.: Key Action Extraction for Learning Analytics. In: Ravenscroft, A., Lindstaedt, S., Delgado Kloos, C., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 320–333. Springer, Heidelberg (2012)
17. Scheffel, M., Niemann, K., Pardo, A., Leony, D., Friedrich, M., Schmidt, K., Wolpers, M., Delgado Kloos, C.: Usage Pattern Recognition in Student Activities. In: Delgado Kloos, C., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 341–355. Springer, Heidelberg (2011)

Designing a Mobile Learning Game to Investigate the Impact of Role-Playing on Helping Behaviour

Birgit Schmitz, Stefaan Ternier, Marco Kalz, Roland Klemke, and Marcus Specht

Center for Learning Science and Technology, Valkenburgerweg 177,
6401 DL Heerlen, The Netherlands
{birgit.schmitz,stefaan.ternier,marco.kalz,roland.klemke,
marcus.specht}@ou.nl

Abstract. Despite research in mobile learning games has intensified over the last decade, there is relatively little research about how individual game mechanisms influence or change behaviour. This article aims at understanding the influence of the game mechanism role-playing and investigates how it can be used to alter behavioural intention. In order to do so, we designed a mobile learning game to train Basic Life Support (BLS) and Cardiopulmonary Resuscitation (CPR). With the game we aim at improving laymen's willingness to help in case of emergency. First, we illuminate the strand of research related to role-playing in the context of BLS and bystander CPR. Second, we describe the pedagogical framework of the mobile learning game that was designed to train BLS and introduce the game engine this development is based on. Third, we present the results from a first prototype testing, which we carried out with medical professionals as well as laymen in order to test game-play usability and interface. This article concludes by outlining the experimental setting of an upcoming study, which will use the mobile learning game to evaluate the influence of the game mechanism role-playing on the willingness to provide bystander CPR in case of emergency.

Keywords: games for health, mobile learning, serious games, role-playing, behavioural intention.

1 Background

In industrialized nations, out-of-hospital cardiac arrest (OHCA) is one of the main causes of death [4]. Although it is known that early initiation of bystander cardiopulmonary resuscitation (CPR) significantly improves the chances of surviving an OHCA [23], one of the main problems still is the availability of trained layman rescuers [12] and low bystander CPR rates [42]. In the 1960s, this led to an extensive introduction of CPR training measures with the intention to increase confidence and the willingness to perform bystander CPR [8]. Up until now, activities in this field have strongly focused on the level of resuscitation skills, teaching methods, frequency of updates or contents of sessions. Despite these activities, the rate of bystander CPR at cardiac arrests has remained low (less than 20%) [34][42]. However important the

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 357–370, 2013.

level of resuscitation skills at performing basic life support, it seems that training on how to provide cardiopulmonary resuscitation is not the only decisive factor.

Studies found that besides the skills and knowledge related to cardiopulmonary resuscitation, other factors such as the perceived risk of infection with a communicable disease during CPR, age or disagreeable physical characteristics, e.g. the presence of blood, influence and even prevent bystanders' willingness to provide CPR [8][23]. Besides, over the past decade studies have increasingly started to investigate the impact of psychosocial factors on bystander initiated CPR [8][10][15][23][25][35].

As Coons reports in his study, 'the relative importance of the reasons for not performing CPR is informative' [10, p. 334]. He points out the potential to change CPR-related attitudes and beliefs and implies different forms of educational intervention to achieve this [10]. Dwyer concedes that in order to optimize the success of CPR training and to modify behaviour, educational interventions should address participant behavioural intentions [15]. The emerging field of games for health too acknowledges this. Games for health promote health related behavioural changes [33][19]. Axelsson too argues that CPR trainings should provide for appropriate models to produce the feelings of personal responsibility and courage required to intervene and to prepare the bystander emotionally for dealing with unexpected and unwanted situations [2].

With this paper we argue that the provision for role-playing into CPR trainings enables such models. The objective of this article is to sample the use of role-playing scenarios for training measures of basic life support (BLS) and bystander CPR and to propose a model that assesses the impact of role-play on helping behaviour. The results could provide valuable information, which might inform future design decisions for BLS and CPR training measures.

2 Related Work

The role-based learning game presented in this paper builds upon a large body of previous research. In the following we outline the main pillars of the game design, i.e. the concept of role-playing in health-care-education and game making environments.

2.1 Role-Playing in Health-Care Education

Kidron [28], in his study on the effectiveness of role-playing exercises, cites Solem [38] who outlined defining goals and characteristics of role-playing. According to his definition role-playing 'deals with emotional and attitudinal antecedents of behaviour in an experiential frame of reference, emphasizes the importance of feelings as sources of behaviour and deals with participants who are placed psychologically inside the problem situation' [28, p. 491].

Since the early nineties, experiential learning has raised great interest amongst health-care-professionals. Experiential learning includes learning through and from experience [11]. It is characterized by learning through doing, role-playing and simulation, all elements thus far shown by literature to enhance CPR training [27]. Kidd and Kendall in their review of *Effective Advanced Cardiac Life Support*

Training [27] refer to the findings of Scherer et al. [37] and Burton [6] who showed that by providing learning scenarios, which include role-playing, learners are more prepared for the unpleasant realities of resuscitation attempts where their CPR skills would be called upon. 'The fact that a resuscitation incident may involve distressed relatives; coupled with increasing evidence advocating the presence of family during resuscitation attempts; it would not be unreasonable to have actors or role-playing learners taking the part of relatives in the simulation' [27, p. 62].

Chamberlain and Hazinski in their article on *Education in Resuscitation* state that 'repeated practice in realistic role-playing scenarios with situations and environments students are most likely to encounter' [7, p. 2578] can increase confidence and the willingness to respond to an emergency. Also, Leigh concluded that 'by participating in simulation scenarios, students can learn to control feelings of panic and their fear of emergency situations' [29, p. 8].

2.2 Game Making Environments

Frameworks for the development of mobile games are currently available. Some explicitly provide for the set up of mobile learning games. Recent studies report on place-based games for language learning or location-based museum games for example that were developed on the base of platforms such as the Myst pervasive game platform [39] or ARIS [22]. Platforms such as 7scenes.com or SCVNGR [26] enable the design, development and deployment of interactive location-based mobile games and experiences. With the aid of such tools teachers are more empowered in utilizing cell phones and creating low-threshold learning opportunities [26]. The MAGICAL project [31] provides a comprehensive overview of state-of-the-art digital game making environments for both desktop and mobile applications.

State-of-the-art game making environments regularly provide for both, the iPhone and the Android Operating System. Working with these systems enables students to bring their own devices to school (BYOD). Many existing platforms are open-source. They usually focus on the display of graphic objects or the handling of touch gestures and frequently embed GPS and QR codes. Few of the reviewed game making environments provide a multiplayer feature based on team play and/or roles though.

In the following chapter we describe the pedagogical framework that informed the game design of *HeartRun* and introduce the game engine ARLearn, which almost comprehensively reflects the required features (e.g. role-based game environment).

3 Pedagogical Framework and Game Design

As prior research has shown, the integration of different roles can bring about a discussion of the psychological and most likely unexpected aspects of cardiac arrest, which may affect the responsiveness of bystanders positively. With the mobile learning game *HeartRun,* we aim at providing a training scenario that offers a framework to test these findings. We took the decision to design a mobile learning game because we wanted to have a scenario that

a) is authentic, i.e. prepares learners to react adequately in a closely related situation. Within the game just as in reality, notification systems are the base for first responders. Dispatch centres send them in case of emergency,
b) includes different locations and different roles. Within the game just as in reality, different roles are involved in case of emergency. While one person gets the nearest automated external defibrillator (AED), another person already starts providing CPR to help the victim. A mobile game can include different locations and different roles.

The game is an additional training measure within the context of a first-aid course or a dedicated basic-life-support training [24]. In the course of the mobile learning game *HeartRun,* we realized the concept of role-playing by providing opportunity to act out the different roles involved in a real case of emergency. In cardiac arrest, it is important to intervene immediately to save seconds and minutes, and to give the most appropriate help possible. Comparable to an unexpected emergency, *HeartRun* involves instant decisions on what to do and the recall of CPR knowledge under unexpected circumstances involving time pressure and stress. This way we intend to enhance psychological preparedness of the rescuer and thus achieving a more prompt and appropriate response. Social psychology presumes that 'the more realistic the situation is made, the closer it is to reality, the better able the subject should be to imagine what he would actually do if he were in the real situation' [18, p. 108].

The *HeartRun* setting comprises an introduction phase, a mobile gaming phase and a debriefing phase, to reflect and share the game experience and to turn it into learning [13].

a) *Introduction Phase*. Players are presented a short introduction to the game, e.g. how to read QR codes with a telephone. They will then be provided with telephones and the game phase will start immediately.

b) *Game Phase*. Students play in teams of two. Every team player is randomly assigned to one of the roles (AED support or bystander). When opening the game, the first message shows, which already relates to the role.

According to the operating mode of ARLearn, opening and reading a message automatically triggers a new message. Messages have different types: text messages, videos, audio messages and questions (single-choice/multiple-choice).

The set-up of the individual messages and the corresponding learning content is related to the *Chain of Survival,* i.e. (a) to prevent cardiac arrest, (b) to buy time, (c) to restart the heart and (d) to restore quality of life. The line of action described there is reflected in the sequence of messages, e.g. making sure that the victim and any bystanders are safe, checking the victim for a response by gently shaking his/her shoulders and asking loudly: "Are you all right?"

Time and location play an important role in the scenario. While Player A (AED support) heads for the AED, player B (bystander) runs to the victim to provide CPR. At the scene of emergency a manikin is provided with which player B interacts. Meanwhile, player A searches the next AED. As soon as he found it and scanned the

QR code attached to it, the game requires him/her to bring the AED to the victim. At the scene of emergency another QR code is placed, which payer A has to scan. This synchronizes the players. Henceforth, both roles get the same information on how to correctly apply the AED. Both players follow the instructions on the screen and apply the AED to the manikin until the ambulance arrives (last message).

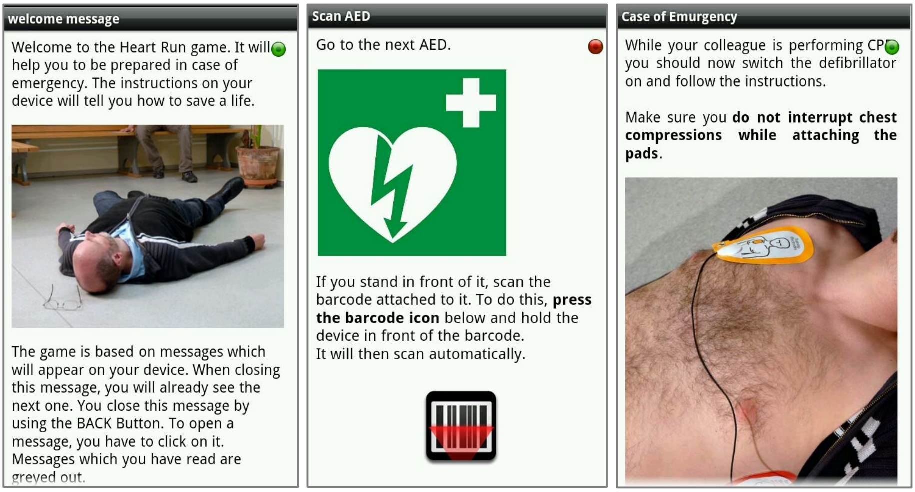


Fig. 1. Game screens *HeartRun*: first message, task, and instruction screen

Also, in the course of the game phase, the player gets an unexpected ALERT message that requires immediate action. An agitated caller pleads: “Quickly, quickly, come to the canteen. George has collapsed. I need your help, please”. This time, no additional information on how to proceed is provided. Then, the game immediately starts to count down 10 minutes and eventually ends.

This part of the game has a third role integrated, the documenter. Immediately after the alert message has arrived, every player immediately needs to start action: find the next AED and get it (role of AED support), go somewhere and provide BLS (role of bystander), record the scene of action with the integrated camera of the smart phone and upload it to the system (role of documenter). After ten minutes the game automatically ends sending the message that the ambulance team has arrived and will take over the care of the victim. The countdown is integrated in order to raise awareness for the span of time, which is given to actually recall actions in relation to BLS.

In the course of the game students upload photos, audio recordings and video sequences, which are considered and evaluated in the subsequent debriefing phase.

c) *Debriefing Phase*. In this last game phase, students meet for debriefing. Within the debriefing phase, they revise and share the knowledge they acquired in the course of the gaming phase. To do so, their recording as well as ideal type of action is presented (gold-standard video). Learners are then required to compare both versions and reflect on things to improve.

HeartRun can be played several times, with participants switching roles. This way, the game allows students to perceive the emergency situation from different perspectives. By ''putting oneself into other's shoes'' they have the chance to experience and control both feelings of panic and fear of emergency [29].

4 The Game Engine

HeartRun was designed with the ARLearn authoring tool, an open source tool suite for educators and learners [39]. ARLearn features a client/server architecture, which enables the creation and management of a) games (reusable instructional design) and b) runs (game instantiations with real time communication) [39]. The ARLearn android client allows playing a run with mobile users. By building on the Google App Engine (GAE) stack, the server architecture offers a scalable web service for content and notification management as well as game state persistence [39].

Games based on ARLearn have a reusable game logic description that can be instantiated in numerous game-runs. Within a game, an author defines items, dependencies between items, game score rules and progress rules. A run defines users grouped in teams. ARLearn enables game-designers to bind a number of content items and task structures to locations and to use game-logic and dependencies to initiate further tasks and activities. While playing, users generate actions (e.g., "read message", "answered question") and responses.

The object model in ARLearn is based on media items, which can be used for messages that users receive at a specified point in time or in relation to a defined event. Further specializations of these media items allow to ask questions (MultipleChoice) or to include multi media (Audio- and VideoObjects).

Items can define dependsOn and disappearsOn conditions for the state transitions. A simple dependency mechanism is put in place to support these conditions:

- Action-based dependencies are triggered by specified actions.
- Time based dependencies bind time offsets to other dependencies.
- Boolean dependencies allow combining other dependencies to logically.

HeartRun uses most of the functionalities the authoring tool provides. The multi-user feature enabled the provision of different player roles and allowed us to send individualized information. The location-based feature for example enabled us to create a realistic game-play setting that supports team play and allows for mixtures of competitive and collaborative games.

Because ARLearn supports the Android operating system (and soon also iOS), the game can be played on commonly used smartphones, which will simplify game distribution for the upcoming larger study. Prior to the forthcoming study we conducted a first prototype testing in order to identify problems that could be addressed beforehand. In the following chapter we present the results from this first prototype testing.

5 Prototype Testing

We tested a first version of the game to assess (a) the *HeartRun* interface and game-play usability (i.e. the integrated features of the prototype) and (b) the technical functionality. Our main focus was on usability aspects such as adequacy of game features in relation to the integrated content (*Chain of Survival*). Therefore, we asked project staff to test-play the game. Four medical experts and three members who are concerned with project related tasks (graphical designer, project manager, and developer) volunteered for the testing. None of them had played *HeartRun* before. The testing was conducted on March 21st 2013 at the general consortia meeting of the EMuRgency project [16] in Leuven.

5.1 Method

We collected data by using a questionnaire and informal interviews after the game session. The questionnaire was based on the System Usability Scale (SUS). We chose the SUS because it is an accepted measure for attitudes toward system usability [30]. It is commonly used in a variety of research projects [5]. The Overall SUS provides a generic questionnaire of ten items, with odd-numbered items worded positively and even-numbered items worded negatively. The items had to be rated on a 5-point Likert-scale. We used the standard Overall SUS score, instead of the 8-item Usability scale, because the Overall SUS allows estimating perceived Learnability along with a cleaner estimate of perceived Usability [30]. Accordingly, we decomposed the score into its Usability and Learnability components. In order to further specify the feedback, we added two additional free text questions: 'For which target group would you use the game?' and 'What would you suggest to improve the game?'

For the gaming session, we equipped the participants with a HTC Desire mobile phone each and randomly assigned the players to one of the two groups (AED and CPR only). Before the testing, the game was installed on the devices and SIM cards were inserted in order to set up an online connection. Prior to the game testing we briefly explained the aim of the game and gave basic instructions on how to use the device. Then the players were asked to start the game. Immediately the first message appeared and people started to play. While playing, intervention was kept to a minimum. The researcher's role during the case study was participant observer. The data we collected were analysed with regard to game usability, learnability, technology use and students' participation.

5.2 Results

The technical game testing delivered valuable feedback with regard to interface and gameplay usability, planned use of the game, integrated features and technical functioning thus safeguarding the upcoming larger study on the impact of role-play on learners' intention to help.

With regard to interface and game-play usability, the *HeartRun* questionnaire for the Overall SUS revealed a mean score of 73,57 with a median of 72,5 and a range from 62,5 to 85. According to standardized interpretation, a SUS score above a 68 is considered above average and on an adjective rating scale could be described as good to excellent [3]. The adjective rating scale established by Bangor, Kortum, and Miller matches the SUS scale and helps to provide a subjective label for an individual study's mean SUS score [3].

In accordance with the proposition by Lewis and Sauro [30], we decomposed the Overall SUS score with the two questionnaire items number four and ten contributing to the Learnability scale and the remaining items contributing to the Usability scale. For perceived Learnability, i.e. the ease of getting used to the application, the questionnaire revealed a mean score of 16,43 with a median of 15 and a range from 12,5 to 20. With regard to perceived Usability, the resulting score ranged from 47,5 to 65, with a mean score of 57,14 and a median of 55.

From the questionnaire it shows that users accepted the game and were motivated to use it in a training context. In the interviews participants further specified that they think the game does not replace traditional BLS training but ideally complements it. Especially the fact that players have to quickly decide what to do in a stressful situation was rated positive feature.

With regard to the target group for the upcoming larger study, participants acknowledged using the game for school children at the age of 10-16 years and young laymen (up to 20 years) as well as persons who are new to a certain environment. According to suggestions from medical experts, we fit some minor wording errors.

Considering system usability and game design *HeartRun* needs adaptation though. Especially the game usability needs some re-work according to the users. They suggested adding further assistance with the device, i.e. directions where to go on the phone and a voice-over function was mentioned. In addition, participants suggested integrating further information on the AED, e.g. an AED movie.

From a technical point of view, the game ran stable. Messages showed sequentially and the two roles involved in the testing were both able to upload audio and video files. We tested the game with two players per team (one player per role for AED support and one for bystander). After the first run, players changed roles and were thus able to test both roles of the game.

We neither tested the third role of the game (documenter) nor the alert function due to time restrictions. This will be part of a second prototype testing, which is due to follow. This second testing will comprise the complete course of action that applies to the experimental setting of an upcoming study. This study will use the game to evaluate the influencing factors of the willingness to provide bystander CPR in case of emergency. The study design we describe in the following chapter.

6 Future Work

We will use the mobile learning game *HeartRun* to evaluate the influence of role-playing on behavioural intention. Because it is difficult to measure the actual behaviour (bystander CPR), intention or willingness to help is a crucial aspect when looking at bystander CPR in case of emergency [8][23].

Over the past decade, a significant body of research has focused on identifying various factors that influence helping behaviour [8][10][15][23][25]. Among this research, the theory of planned behaviour (TPB), introduced by Ajzen [1], has received attention. First research approaches in the field indicate that this theory should be capable of explaining bystander's motivations with regard to CPR training and performance [15][41]. Ajzen's theory implies that behavioural intention is the most influential predictor of behaviour and uses intention as a proximal measure of behaviour [1]. It suggests that behavioural intention can be induced through modification of one or more of three antecedents of intention: attitude toward the behaviour, subjective norm and perceived behavioural control. The TPB 'proposes that the strength of an individual's intention (or motivation) to engage in a behaviour, and the degree of control they feel they have over that behaviour (perceived behavioural control) are the proximal determinants of engaging in it' [41].

Findings on attitude and self-efficacy (perceived behavioural control) in social psychology indicate that role-playing influences a person's behaviour. Prior research investigated the effect of a serious game on role-taking and willingness to help [33]. Peng, Lee and Heeter aimed at understanding how interactive digital games affect role-playing and helping behaviours. They found that the presentation mode (interactive game) positively influenced participants' willingness to help and that role-taking partially mediated this relationship [33]. They state that during the role-taking process, an individual goes beyond his or her typically egocentric means of perceiving the world to contemplate a different point of view. Also Pavey, Greitemeyer and Sparks resume that role-taking is likely to produce empathy, which in turn has been found to be a strong predictor of helping behaviour [32]. Chamberlain and Hazinski claim that realistic role-playing scenarios can increase confidence and the willingness to respond to an emergency [7].

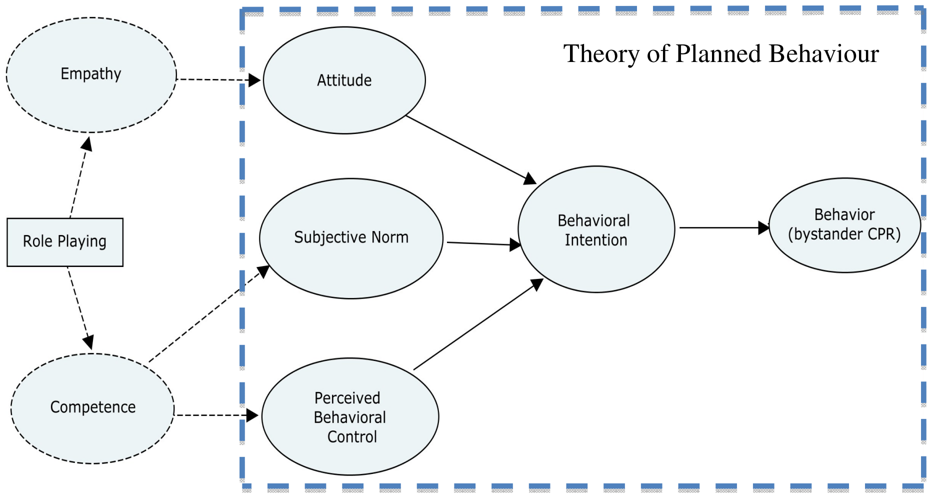


Fig. 2. Proposed research model

An important factor related to CPR intervention is competence. BLS and CPR training increases laypersons' confidence and willingness to perform bystander CPR on a stranger [8]. Even doctors who are competent in performing resuscitation may fail to apply their skills, unless they have a strong belief in their own capabilities [9].

Drawing upon findings in game studies, social psychology and medical education our research extends the Theory of Planned Behaviour by introducing variables of empathy, competence and role-playing to TPB in order to predict helping behaviour. We hypothesize that empathy, competence and role-playing influence the willingness to help in case of emergency.

Figure 2 represents the proposed research model. It incorporates the three additional constructs empathy, competence and role-playing, which we assume to influence the existing variables if the TPB (attitude, subjective norm and perceived behavioural control). To explain the constructs of the model and the related hypothesis would go beyond the scope of this paper. This will be part of a subsequent article, which will report the experimental settings and results.

A connatural approach was introduced by Vaillancourt et al. [41] in order to identify barriers and facilitators of CPR training and CPR performance. Their work did not consider the concept of role-playing as an independent variable for the model.

6.1 Method

To test the hypotheses, we use a true experiment with students from a secondary school. Participants are between 12 and 16 years old. The evaluation setting will include:

a) Intervention (BLS training measure including the game) directed at changing behavioural intention
b) Post-test of specific and generalized value of the concepts described and behavioural intention

The study will not include a pre-questionnaire. We assume that a pre-test would indicate the concepts and conceptual interrelations we are about to investigate. Post-test data will be collected immediately after the training.

The findings of Plant and Taylor support this setting. They reviewed literature to identify methods of CPR, AED and first aid training children have been successful. Results suggest that the use of CBT, 'virtual worlds' and 'multiplayer online simulation' in CPR training 'could be an attractive training and/or retention tool to use in this age group' [34, p. 3].

6.2 Measures

The constructs in the model are measured with standard items adapted from previous research. All questionnaire items use a 7-point Likert-type scale where 1=completely disagree and 7=completely agree. In order to evaluate behavioural intention, we will provide a hypothetical scenario to the students. The scenario includes a short passage of information about a situation, which is likely to occur [25]. TPB items will be adapted from Francis et al. [17]. A checklist with predefined criteria will assess proficiency in CPR and AED skills. We will measure empathy with items adapted from the four empathy scales by Davis [14].

6.3 Data Analysis

In order to corroborate data conformity between the empirical data and the presumed model, we will first conduct a confirmatory factor analysis (CFA). We will then use structural equation modelling (SEM) techniques to examine the causal model. Structural Equation Modelling is commonly used in predicting and explaining health behaviour [36].

The proposed research model and hypothesis testing will be carried out using variance-based partial least squares (PLS) path modelling. This approach to structural equation modelling (SEM) techniques, as opposed to the covariance-based SEM, is a 'prediction-oriented variance-based approach that focuses on endogenous target constructs in the model and aims at maximizing their explained variance' [20, p. 312]. Recent review results of PLS-SEM applications [21] suggest that PLS-SEM allows the use of multivariate analysis tool for small sample size (and is even advantageous), provided the fact that the fundamentals of sampling theory are considered. As the planned experiment is rather complex in procedure and with large sample sizes, this is an important factor to consider.

7 Conclusion

This paper's contribution to research in the field is based on four different pillars. Firstly, we provided an overview of related work on role-playing with regard to BLS and bystander CPR. Secondly, we exemplary described the educational framework of the mobile learning game *HeartRun.* Prior research on behavioural issues in BLS and CPR training informed the game design. Subsequently we depicted the technical base, the game engine ARLearn. Thirdly, we provided results from a first prototype testing, which was carried out with medical professionals as well as laymen in order to test game-play usability and interface. Eventually, this article outlined a new approach within the field of research on resuscitation. The experimental setting of a forthcoming study will use the mobile learning game, to better understand the influence of role-playing on the willingness to provide bystander CPR in case of emergency. To do so, it applies the Theory of Planned Behaviour [1]. Data from the experiment will be analysed using structural equation modelling.

Acknowledgement. This publication was partly financed by the European Regional Development Fund (ERDF), regions of the Euregio Meuse-Rhine and the participating institutions under the INTERREG IVa program (EMR.INT4-1.2.-2011-04/070, http://www.emurgency.eu).

References

1. Ajzen, I.: The theory of planned behaviour. Organizational Behaviour and Human Decision Process 50, 179–211 (1991)
2. Axelsson, A., Herlitz, J., Fridlund, B.: How bystanders perceive their cardiopulmonary resuscitation intervention; a qualitative study. Resuscitation 47(1), 71–81 (2000)
3. Bangor, A., Kortum, M., Miller, J.: Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. Journal of Usability Studies 4(3), 114–123 (2009)
4. Breckwoldt, J., Schlösser, S., Arntz, H.-R.: Perceptions of collapse and assessment of cardiac arrest by bystanders of out-of-hospital cardiac arrest (OOHCA). Resuscitation 80(10), 1108–1113 (2009)
5. Brooke, J.: SUS - a quick and dirty usability scale. In: Usability Evaluation in Industry, pp. 189–194. Taylor & Francis, London (1996)
6. Burton, B.C.: Training done right. Occupational Health and Safety 71, 74 (2002)
7. Chamberlain, D.A., Hazinski, M.F.: Education in Resuscitation. In: An ILCOR Symposium Utstein Abbey Stavanger, Norway, June 22-24 (2001), `http://circ.ahajournals.org/content/108/20/2575`, doi:10.1161/01.CIR.0000099898.11954.3B (retrieved November 12, 2012)
8. Cho, G.C., Sohn, Y.D., Kang, K.H., Lee, W.W., Lim, K.S., Kim, W., Oh, B.J., Choi, D.H., Yeom, S.R., Lim, H.: The effect of basic life support education on laypersons' willingness in performing bystander hands only cardiopulmonary resuscitation. Resuscitation 81(6), 691–694 (2010)
9. Coolen, E.H.A.J., Loeffen, J.L., Draaisma, J.M.T.: Enhancing self-efficacy for paediatric resuscitation skills in the undergraduate medical curriculum. Resuscitation 81, 131–132 (2009)
10. Coons, S.J., Guy, M.C.: Performing bystander CPR for sudden cardiac arrest: Behavioural intentions among the general adult population in Arizona. Resuscitation 80, 334–340 (2009)
11. Cooper, S., Libby, J.: A review of educational issues in resuscitation training. J. Clin. Nursing 6, 5–10 (1997)
12. Creutzfeldt, J., Hedman, L., Felländer-Tsai, L.: Effects of pre-training using serious game technology on CPR performance – an exploratory quasi-experimental transfer study. Scand. J. Trauma Resusc. Emerg. Med. 20(79) (2012), doi:10.1186/1757-7241-20-79
13. Crookall, D.: Serious Games, Debriefing, and Simulation/Gaming as a Discipline. Simulation & Gaming 41(6), 898–920 (2011)
14. Davis, M.H.: A Multidimensional Approach to Individual Differences in Empathy. JSAS Catalogue of Selected Documents in Psychology 10, 85 (1980)
15. Dwyer, T., Mosel Williams, L.: Nurses' behaviour regarding CPR and the theories of reasoned action and planned behaviour. Resuscitation 52(1), 85–90 (2002)
16. EMuRgency: New approaches for resuscitation support and training, project website: `http://www.emurgency.eu`. (last visited March 22, 2013)
17. Francis, J.J., Eccles, M.P., Johnston, M., Walker, A., Grimshaw, J., Foy, R., Kaner, E.F.S., Smith, L., Bonetti, D.: Constructing questionnaires based on the Theory of Planned Behaviour. Centre for Health Service Research, University of Newcastle, Newcastle upon Tyne (2004)
18. Freedman, J.L.: Role Playing: Psychology by Consensus. Journal of Personality and Social Psychology 13(2), 107–114 (1969)
19. Games for Health. Exploring the intersection of videogames + health, `http://gamesforhealth.org/` (last visited February 18, 2013)

20. Hair, J.F., Ringle, C.M., Sarstedt, M.: Partial Least Squares: The Better Approach to Structural Equation Modelling? Long Range Planning 45(5-6), 312–319 (2012)
21. Hair, J.F., Ringle, C.M., Sarstedt, M.: Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance. Long Range Planning 46, 1–12 (2013)
22. Holden, C.L., Sykes, J.M.: Leveraging Mobile Games for Place-Based Language Learning. International Journal of Game-Based Learning 1(2), 1–18 (2011)
23. Johnston, T.C., Clark, M.J., Dingle, G.A., FitzGerald, G.: Factors influencing Queenslanders' willingness to perform bystander cardiopulmonary resuscitation. Resuscitation 56(1), 67–75 (2003)
24. Kalz, M., Schmitz, B., Biermann, H., Klemke, R., Ternier, S., Specht, M.: Design of a game-based pre-hospital resuscitation training for first responders. Paper Accepted for SouthCHI Conference, Maribor, Slovenia, July 01-03 (2013)
25. Kanstad, B.K., Nilsen, S.A., Fredriksen, K.: CPR knowledge and attitude to performing bystander CPR among secondary school students in Norway. Resuscitation 82, 1053–1059 (2011)
26. Kelsey, M., Murphy, W.: Engaging Students and Improving Teacher Practices Through Simple 21st Century Technology: Cell Phones. In: Resta, P. (ed.) Proceedings of Society for Information Technology & Teacher Education International Conference, pp. 1876–1879. AACE, Chesapeake (2012)
27. Kidd, T., Kendall, S.: Review of effective advanced cardiac life support training using experiential learning. Journal of Clinical Nursing 16(1), 58–66 (2007)
28. Kidron, G.A.: The Effectiveness of Experiential Methods in Training and Education: The Case of Role Playing. The Academy of Management Review 2(3), 490–495 (1977)
29. Leigh, G.T.: High-fidelity patient simulation and nursing students' self-efficacy: A review of the literature. International Journal of Nursing Education Scholarship 5(1), 1–17 (2008), doi:10.2202/1548-923X.1613
30. Lewis, J.R., Sauro, J.: The Factor Structure of the System Usability Scale. In: Kurosu, M. (ed.) HCD 2009. LNCS, vol. 5619, pp. 94–103. Springer, Heidelberg (2009)
31. MAGICAL: Making Games in Collaboration for Learning, project website: `http://www.magical-project.net` (last visited June 22, 2013)
32. Pavey, L., Greitemeyer, T., Sparks, P.: I help because I want to, not because you tell me to: Empathy increases autonomously motivated helping. Personality and Social Psychology Bulletin 38(5), 681–689 (2012)
33. Peng, W., Liu, M.: An overview of using electronic games for health purposes. In: Ferdig, R. (ed.) Handbook of Research on Effective Electronic Gaming in Education, pp. 388–401. IGI Global, Hershey (2008)
34. Plant, N., Taylor, K.: How best to teach CPR to schoolchildren: A systematic review 84(4), 415–421 (2013)
35. Query, P.: Attitudes towards the willingness to help. Journal of Undergraduate Psychological Research 1, 18–21 (2006)
36. Rasmussen, M., Krolner, R., Klepp, K.I., Lytle, L., Brug, J., Bere, E., Due, P.: Determinants of fruit and vegetable consumption among children and adolescents: A review of the literature. Part I: quantitative studies. International Journal of Behavioural Nutrition and Physical Activity 3(1) (2006), doi:10.1186/1479-5868-3-22
37. Scherer, Y., Bruce, S., Graves, B., Erdley, W.S.: Acute care nurse practitioner education: Enhancing performance through the use of clinical simulation. Advanced Practice in Acute Critical Care 14, 331–341 (2003)

38. Solem, A.: Human Relations Training: A Comparison of Case Study. Personnel Administration 23, 29–37 (1960)
39. Laine, T.H., Vinni, M., Sedano, C.I., Joy, M.: On designing a pervasive mobile learning platform. ALT-J, Research in Learning Technology 18(1), 3–17 (2010)
40. Ternier, S., Tabuenca, B., Specht, M.: ARLearn: Learning activities and interaction in augmented reality. In: Specht, M., Multisilta, J., Sharples, M. (eds.) Proceedings of the Mobile Augmented Reality for Education Workshop, Helsinki, Finland, October 16-17, pp. 10–13 (2012)
41. Vaillancourt, C., Grimshaw, J., Brehaut, J.C., Osmond, M., Charette, M.L., Wells, G.A., Stiell, I.G.: A survey of attitudes and factors associated with successful cardiopulmonary resuscitation (CPR) knowledge transfer in and older population most likely to witness cardiac arrest: design and methodology (2008), http://www.biomedcentral.com/1471-227/X/8/13 (retrieved February 25, 2013)
42. Vaillancourt, C., Wells, G.A., Stiell, I.G.: Understanding and improving low bystander CPR rates: A systematic review of the literature CJEM 01/2008 (2008)

Designing MOOCs for the Support of Multiple Learning Styles

Franka Grünewald, Christoph Meinel,
Michael Totschnig, and Christian Willems

Hasso Plattner Institute, University of Potsdam
Prof.-Dr.-Helmert-Str. 2-3, Potsdam, Germany
{firstname.lastname}@hpi.uni-potsdam.de

Abstract. "Internetworking with TCP/IP" is a Massive Open Online Course, held in German at openHPI end of 2012, that attracted a large audience that has not been in contact with higher education before. The course followed the xMOOC model based on a well-defined sequence of learning content, mainly video lectures and interactive self-tests, and with heavy reliance on social collaboration features. From 2726 active participants, 38% have participated in a survey at the end of the course. This paper presents an analysis of the survey responses with respect to the following questions: 1) How can a MOOC accommodate different learning styles and 2) What recommendations for the design and organization of a MOOC can be concluded from the responses? We finally give an outlook on challenges for the further development of openHPI. Those challenges are based on didactical and technical affordances for a better support of the different learning styles. We propose an evolution of the xMOOC, that bridges the gap to the cMOOC model by developing tools that allow users to create diverging paths through the learning material, involve the user personally in the problem domain with (group) hands-on exercises and reward user contributions by means of gamification.

Keywords: massive open online courses, learning styles, culture of participation, active learning, gamification.

1 Introduction

The concept and format of "Massive Open Online Courses" (MOOC) have been invested with hopes for radical changes in higher education, due to their potential to make high quality teaching accessible to everyone with broadband Internet access and motivated to invest their time into concentrated learning. While all MOOCs share the goal of bringing together thousands of learners into a common event, they largely differ with respect to the underlying concept of openness [11]. Minimally MOOCs are open because access is not restricted by cost, affiliation, or any other type of privilege. They differ with respect to the openness of the learning content and the learning process. Siemens [14] has suggested to distinguish between xMOOCs which rely on the traditional lecture format supplemented with interactive exercises and discussion boards, and cMOOCs that

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 371–382, 2013.

are based on a connectivist pedagogy [13] that invites learners to engage in a self-organized and social learning process. While xMOOCs tend to use learning materials with proprietary licenses and a relatively closed, predefined schedule, cMOOCs often make use of open educational resources, and allow learners to co-construct the learning process through their interactions. It has been argued (T. Bates cited by [11]) that the pedagogy of xMOOCs is better suited for learning domain knowledge that can be mastered through repetitive practice, but that only cMOOCs allow learners to acquire higher order creative skills.

The xMOOC format derives its name from the platform edX[1], founded by Harvard University and the Massachusetts Institute of Technology as joint venture for hosting online courses that extend the reach of the university's teaching to a massive audience. The first experiments with the xMOOC format though are ascribed to courses offered by professors in Stanford university[2] that have later inspired the two commercial projects, Udacity[3] and Coursera[4]. The cMOOC concept was derived from a course experiment led in 2008 by Canadian educational researchers, George Siemens and Stephen Downes, and the discussion it generated about the pedagogical theory of connectivism that conceives learning as the creative and social process of connecting nodes of knowledge [9].

In this paper, we report on our experiments in an xMOOC where we included practical exercises in order to go beyond the purely theoretical presentation of learning content and to invite learners to relate and to apply knowledge to their everyday environment. We believe that this strategy

– will allow mediation of the dichotomy of cMOOC and xMOOC,
– is attractive for different learning styles
– has the potential to nurture an active participatory culture.

openHPI [6] is a platform for xMOOCs, offered by the Hasso Plattner Institute (HPI) in Potsdam , Germany. In November/December 2012 it hosted the course "Internetworking with TCP/IP" (the first xMOOC in German language) that attracted a large audience that has partly not been in contact with higher education before.

The majority of our course participants belong to the 20-29 and 30-39 age groups (each approx. 30%). About 20% belong to the group from 40 to 49 and a remarkable high share of 16% comes from the "silver surfers" group above 50 years. The remaining 4% are pupils of 19 years and younger. The youngest participant stated his age with 12 years, the oldest with 91. About 24% of the participants said, that they never went to university, 21% chose a B.Sc. as their highest degree, 25% the M.Sc. or equivalent and 4% finished a PhD degree. The

[1] https://www.edx.org
[2] Artificial Intelligence by Sebastian Thrun and Peter Norvig (https://www.ai-class.com/, Machine Learning by Andrew Ng (https://www.ml-class.org/) and Introduction to Databases by Jennifer Widom (https://www.db-class.org)
[3] https://udacity.org
[4] https://coursera.org

remaining 26% answered with "other" when asked for their highest degree. Since openHPI focuses on ICT topics, we also ask for our users background in ICT on registration: approx. 6% stated to have no experience, 32% declared themselves as "beginners", 45% as "advanced" and 17% as "experts".

From a survey delivered to all of the approximately 10,000 registered participants (out of which 2,726 participated actively), we obtained more than 1,000 responses, that allow us to understand the motivations, conditions and expectations for taking part in the course, and to obtain a high number of valuable suggestions for improving the course content and format.

The structure of the paper is as follows: First, we explain the concepts of experiential learning and culture of participation. Next, we present the results of the survey and then our implications in form of design guide lines. We conclude with an outlook on further development of the platform.

2 Theory

Experiential learning describes a didactical model, where learning is not primarily based on abstract theorization, but on a holistic cycle that includes concrete experience, reflective observation, abstract conceptualization and active experimentation [8]. These phases are organized in two dimensions: perceiving (from abstract to concrete) and processing (from reflective to active). Learners prefer one of four combinations of these two dimensions:

- **Divergers** combine concrete experience with reflective observation. They learn from examples and can analyze these from different perspectives.
- **Assimilators** combine abstract conceptualization with reflective observation. They prefer learning from theoretical models.
- **Convergers** combine abstract conceptualization with active experimentation. They learn by processing ideas and concentrating on precise problems.
- **Accomodators** combine concrete experience with active experimentation. They learn from experiments and match models to their obtained insights.

xMOOCs predominantly cater to the assimilating learning style through the presentation of concepts and the video format that invites to reflective observation. In order to make the learning process more holistic, MOOCs need to integrate activities that allow active experimentation and that relate to concrete experience.

What is still missing in Kolb's theory is an explicit recognition of the social dimension of learning [17]. The social dimension of learning takes into account that learning occurs together with others in all kinds of social situations or contexts [1]. Wenger argues that a learner needs to participate in a community in order to understand and create meaning [16]. ICT in distance education has the role to to allow students to learn from and support each other despite the physical separation [10]. We believe that the problem ob the missing social dimension in Kolb's theory of learning can be alleviated through the inclusion of Fischer's concept *culture of participation*. This is the case, because Fischer's concept deals

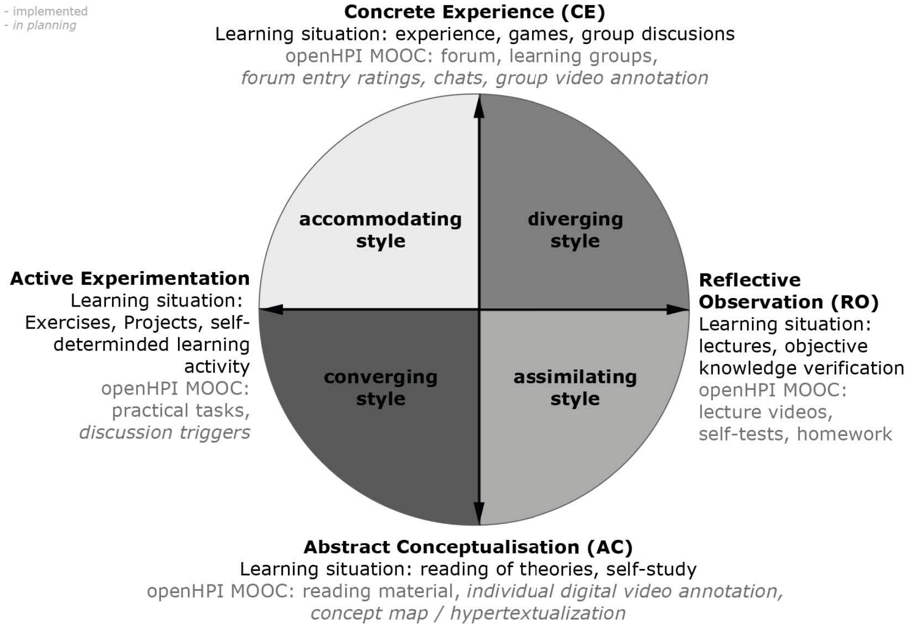


Fig. 1. Learning Styles by to Kolb (according to [8] and [15] pp. 67/70) and their specification in openHPI

with guidelines for socio-technical systems to be implemented in a participative manner so that individual are supported to engage in collaborative activities. And Wheeler explains, that the weekness of Kolb's theory is exactly that, not to consider the interactive digital media of today's world that is socially rich [17].

Gerhard Fischer [5] suggests design guidelines for socio-technical systems aiming at stimulating participation. Through certain key elements within the platform, social exchange and collaboration should be enabled and supported. The culture of participation consists of three parts:

- **meta-design**: where collaborative design is enabled by the infrastructure;
- **social creativity**: that shall support collaboration among learners;
- **different levels of participation** (see figure 2): those levels should allow different degrees of engagement with the system and its content. The different levels range from totally unaware consumers that have no knowledge about the possibilities they have to participate to the meta-designers who even go beyond the boundaries of the given environment and create new knowledge, tools and workflows.

Intrinsic motivation is basis on which the culture of participation is build upon. This motivation can be triggered by group support, the feeling of a common purpose and a collaborative creativity [5]. Fischer's design guidelines include the following outlines:

– Human-Problem Interaction as an advancement over human-computer-in”-ter”-ac”-tion shall be supported with the help of meta-design environments in order to make problem owners responsible and encourage social creativity.
– Underdesign to encourage lively and open information creation and encourage individual workflows instead of limitating the participants by fixed environments, contents and processes.
– Support diverse levels of engagement in order to respect and consider the different motivations, pre-requisites and needs of the participants and to smooth the path to more challenging roles.
– Awareness and rewards for contributions should be supported in order to strengthen the social incentives for participants. Those incentives include the possibility to build up a reputation within the community and receive feedback.
– Parallel development of the community and resources for system development shall be supported by encouraging cross-pollination between those two.

They are underlying some of our design guidelines for MOOCs that will be presented in the next but one section.

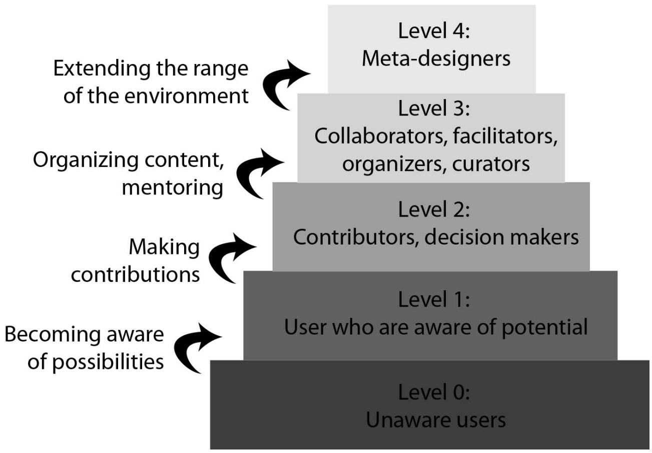


Fig. 2. Ecologies of Participation according to Gerhard Fischer [5]

Dick and Zietz [3] further analyzed different motivations within these cultures of participation. They identified social norm, social proof as well as peer pressure as the main motivational factors. The key enabler for this behavior to be triggered is the awareness of other group members’ actions within the socio-technical-system. Dick and Zietz therefore concluded that instead of trying to make people more active, system designers need to focus more on the awareness of the users for the action they may take and the reward they may get from it.

The next section will deal with a mostly qualitative, but in key points also quantitative analysis of the participants of the “Internetworking” course. We aim at providing further information about which offerings were most accepted and useful as well as which further developments people wish for. We focus our analysis on elements identified within the learning styles and cultures of participation theories.

3 Survey Evaluation

At the beginning of 2013 we conducted a study among the students of the first German speaking MOOC at our institute, that is to our knowledge the German speaking MOOC with the largest number of participants so far. From the 2726 active participants of the course, 42,3% have taken part in the survey. The large number of questionnaires completed (n=1153) allows us to statistically evaluate the results. Nevertheless, there is the possibility of self-selection bias since we did not capture if only very successful participants or those who had a very positive experience completed the questionnaire.

The survey allowed us to confirm a high degree of satisfaction amidst the participants with the course content and structure: High satisfaction with the expertise of the learning material was expressed by 92.1% (70.7% very satisfied, 21.4% rather satisfied) for the lecture videos, 88,8% (63.7% very and 25.1 rather satisfied) for the tutorial videos, 89.8% (58.3% very and 31.5% rather satisfied) for the slides, 70.8% (44.8% very and 26.0 rather satisfied) for the reading material, 87.7% (56.6% very and 31.1% rather satisfied) for the quizzes. At the same time, the usefulness of the different types of learning materials was also confirmed to a large extent, as can be seen in figure 3. Interestingly the social features of the platform (forums, learning groups) were not seen as having a positive impact on the learning success. We also asked users to express their self-assessment of their expertise in five different topic domains before the start and after the end of the course, on a scale with the following values: 0 (no knowledge), 1 little knowledge, 2 substantial knowledge, 3 expert knowledge. The average delta for the cumulative knowledge across the five domains was 4.86 points, i.e. the average user was able to advance approximately one level in each of the five domains. 8.6% of the users even expressed an advancement of 10 points and more. We cannot eliminate the optimism bias, though, since the knowledge delta was calculated based on a self-evaluation of the participants.

The survey also gave us access to qualitative feedback and recommendations for improvements to the platform and to the course content, from which we will deduce guidelines in the following section. In the following, we summarize some salient topics:

- **Consistency** Learners expressed very thoughtfully their dissatisfaction with several occurrences of inconsistencies in the content. For example when quizzes made references to concepts that had not been covered in the preceding video lecture, when definitions were contradictory or not sufficiently precise.

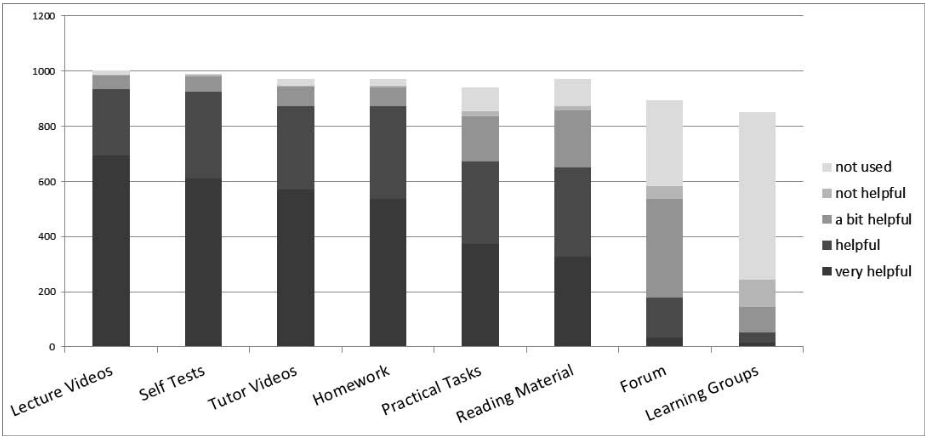


Fig. 3. Reception of the learning materials by the participants of openHPI

- **Multimedia** Participants suggested to move beyond the traditional lecture format and make more intensive use of visualizations, animations and simulations, which helps grasp complex concept relationships.
- **Hypertext** Numerous respondents requested more extensive collections of links that allow them to discover valuable resources on the Internet.
- **Synchronous communication** Several participants expressed the wish for more immediate communication, which can be supported through the establishment of private and group chat functions.
- **Practical relevance** Most important, survey respondents asked for still more practical examples and exercises, for example by using virtual laboratories, or by presenting challenges that would require investigating the learner's own network environment.

4 Design Guidelines for MOOCs to Support Experiential Learning

MOOCs have the potential to deliver high quality learning experiences to an unprecedented high number of learners. Our survey confirmed on the one hand that learners acknowledge the quality of video lectures and textual learning resources, and we conclude that these effectively support the assimilating learning style and that they explain the high level of self-estimated positive learning results. On the other hand, learners ask for more intense support of active experimentation and are interested in relating the concepts to their own experience. MOOCs can go beyond the concentration on the assimilating style through the adoption of the following guidelines, that are based on the integration of a social dimension into Kolb's concept of experiential learning. First we describe how learning tools can target the four learning styles separately, then we show how they can be integrated into a holistic process through the integration of a culture of participation.

- The accommodating learning style should be supported through practical exercises that call upon learners' existing experience. In the "Internetworking" course we included tasks like using Wireshark[5] to find out about certain details of the protocol usage within one's own computer. The hands-on tasks were introduced by a tutor video (or screencast) and heavily questioned and discussed in the forums. Inquiry-based learning is a long-standing paradigm in the educational sciences, but scaling these approaches to the context of a MOOC is still an open challenge.
- The diverging learning style is inspired by the inclusion of many perspectives, and the high number of participants in a MOOC can be leveraged to create these perspectives. In order to nurture a vibrant discussion culture, we introduced discussion triggers, discussion threads where the community should discuss issues strongly related to the course content, but not directly covered. This can even result in learners contributing these triggers, as we could observe with a posting containing an assignment question concerning the behavior of nodes on a network route. Salmon [12] has developed an extensive framework on best practices for stimulating discussions in on-line learning.
- The importance of consistent and well presented learning materials has been mentioned in the preceding chapter, and this is particularly true for learners that prefer an assimilating learning style. Beyond the necessary quality assurance, we also suggest to provide glossaries for quick reference and concept maps that facilitate the cognitive orientation in the knowledge domain. More generally speaking, learning materials in MOOCs should not fall back behind the advancements made in educational hypertexts during the last decades: Hypertextual links should allow learners to understand relations between concepts, and also to navigate beyond the boundaries of the course.
- Converging learners are best inspired by visualizations, simulations and experiments that involve problem solving and decision making. MOOC platforms like Udacity and edX provide virtual environments that allow learners to directly interact with simulations for electronic circuitry or a programming environment. At openHPI, we are actively investigating how virtual IT environments can be made scalable for massive participation. Fischer et al. [4] have shown that presenting learners with content-specific visualization tools can foster the collaborative construction of knowledge.

The massive nature of participation in a MOOC creates new opportunities for strengthening the social dimension of learning. As we have described, a learning community with an active culture of participation can constitute an environment where the four phases of experiential learning are more dynamically intertwined. In the following paragraphs we present in more detail how active participation in a community allows learners to enrich abstract concepts through group annotations, to engage more actively in experiments that allow for creative interpreta-

[5] Wireshark is a well-known tool for the inspection of network traffic – a so-called network sniffer – that is used for network administration and diagnosis as well as for eavesdropping, see `http://www.wireshark.org/` .

tion of concepts, to develop new perspectives on concrete experience from group discussions and to reflect one's learning process more thoroughly by linking it to how one's peers learn.

Different Participation Levels. In order to support social interaction, different levels of engagement should be open for the participants, dependent on their interest in social interaction. The levels can be seen adjacent to the levels proposed by Fischer in his design guidelines for the culture of participation [5].

– Level 0 in the hierarchy of participation levels in MOOCs is describing the passive consumer, who watches videos, consumes the reading material, does his exercises for himself.
– Level 1 is the aware consumer, who is reading the forum, but not actively participating in other activity.
– Level 2 is the active participant who writes forum posts himself and participates in learning groups.
– Level 3 is the enabling participant who starts own forum threads, user groups and chats and triggers discussions and other group learning activities, like manuscript writing in groups.
– Level 4 describes meta-designing participants who move beyond the given learning platform and implement their own games, tools, tutorials, blogs and other material that fellow learners may use.

Awareness, Judgment and Reward for Contributions. As awareness is the key to active participation and social judgment and reward a major motivational factor [7], these should be supported intensively. Rating of forum entries, a public board for the most active users, and promotions to forum moderators are possibilities for the implementation of this design guideline.

These aspects are approached by enhancing the openHPI platform software with a comprehensive set of gamification features. Gamification is defined as "the use of game elements and game design techniques in a non-games context" [2], where the game elements toolbox contains concepts like points and scores, levels, badges, progress bars, leader boards, challenges (or quests), etc. Some common functionality of e-learning platforms fit into a gamification system, in precise all kinds of quizzes (which gain points) or the reception of learning material (which can show up in an overall progress display). For other functionality specific to MOOC platforms – in particular discussion boards – there are sophisticated best practice examples, such as *Stack Overflow* (see `http://stackoverflow.com/`) that show, how gamification can result in an enhanced contribution feedback experience. The *Stack Overflow* discussion boards for example allow voting on everything, i.e. questions, answers or even comments to answers. Users gain points and badges based not only on the number of contributions, but also on the quality of their contributions as experienced by the community. This concept can be mapped directly on most user generated content in the context of a learning platform and will be part of the openHPI gamification feature set.

Stimulate Contributions. In order to evolve from a closed and pre-defined learning setting to an open learning culture, participants should be stimulated to provide own content and enhancements to existing content. openHPI participants for example generated their own audio podcasts to our video lectures and their own exercises – even though the current state of the platform does not provide any facilities for users to integrate these contributions. The participants actually used the course forum to make their contributions available to the learning community [6]. Participants' video tutorials and other enhancements mentioned in participation level 4 already are further contributions that can be thought of. The task of the platform provider is to offer a space where those contributions can be made accessible to fellow learners. Current plans for a platform extension cover, among other things, tools for:

- community-driven subtitles and translations for the lecture videos;
- uploading participants' media files (podcasts, tutorial videos, mind maps, etc.);
- a *battle ground* for user generated quizzes and challenges.

Supplying rating capabilities for all of those items is a suitable way of quality assurance and integrates seamlessly with the above-mentioned gamification concept. Besides this, many participants positively replied to the question about the usefulness of group annotation features for videos. Considering the only moderate acceptance of general learning groups, this new group annotation feature should be closely linked to the lecture video and should again incorporate awareness and reward mechanisms.

Human-Problem Interaction. An approach for leveraging human-problem interaction is the provision of hands-on exercises. There are numerous solutions for courses and online laboratories in the domain of programming (e.g. *Codeacademy* or *CodingBat*[6]), and databases (e.g. Standfords online course *Introduction to Databases*, mentioned before as *DB-Class*). More general approaches for such laboratory environments from the past years usually build on virtualized computer labs that provide remote access to virtual machines running on a central server (respectively the cloud) or are distributed on removable media (rather unsuitable for MOOCs). The problem with these virtual laboratories in the context of a social learning experience is that the assignment tasks usually are static or must be personalized manually, which doesn't scale for a massive amount of users. Thus, these hands-on assignments would not be cheating proof, since learners just could share the results of a practical task allowing other users to omit the task but still being able to solve the assignment. The authors of [18] propose a more generic approach for the automatic assessment of hands-on exercise assignments: the lab management system asserts a student- and task-specific pre-condition that is configured inside a training machine before the student can get access. During the exercise, the student can reveal a "secret" that is affected by the pre-condition and thus prove the successful completion of

[6] see `http://www.codeacademy.com` and `http://codingbat.com`

the practical task by submitting the unique secret value to a quiz environment. Future openHPI courses will integrate a virtual lab automatic online assessment.

5 Conclusion and Future Work

Pedagocical models and technical frameworks for MOOCs are in the focus of scholarly debate and media attention. Based on the experience of openHPI's "Internetworking" course and on an evaluative survey, we have presented arguments for a future development of the xMOOC model that bridges the gap towards the cMOOC model: 1) Learning materials could be enriched through concept maps and hypertextual links that allow diverging, learner-defined paths; 2) Hands-on exercises allow learners to feel personally involved in the problem domain through their active experimentation and to grasp the complex relations to their own concrete experience; 3) Group discussions that support awareness, and reward contributions, allow learners to feel responsible and to collaboratively strengthen the learning process and to provide richer perspectives for reflective observation. This development heavily depends on the emergence of a culture of participation, where learners are motivated to contribute to a network of resources. We have outlined some guidelines for supporting this culture, that we are adopting and improving in our current research and development activities: human-problem interaction in scalable virtual laboratories, and learning services and practical tasks that connect with learners' living environment; gamification features that increase the learning motivation and create responsibility and engagement; social communication tools that allow users to evolve from lower to higher degrees of participation.

References

1. Bliss, J., Light, P., Shaljho, R.: Learning Sites: Social and Technological Resources for Learning. Elsevier Science Inc., New York (1999)
2. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011, Tampere, Finland, pp. 9–15. ACM (2011)
3. Dick, H., Zietz, J.: Cultures of Participation als eine Persuasive Technologie. I-com 9(2), 9–15 (2011)
4. Fischer, F., Bruhn, J., Gräsel, C., Mandl, H.: Fostering collaborative knowledge construction with visualization tools. Learning and Instruction 12(2) (2002)
5. Fischer, G.: Understanding, Fostering, and Supporting Cultures of Participation. Interactions 80(3), 42–53 (2011)
6. Grünewald, F., Mazandarani, E., Meinel, C., Teusner, R., Totschnig, M., Willems, C.: openHPI - a Case-Study on the Emergence of two Learning Communities. In: IEEE Global Engineering Education Conference (2013)
7. Kimmerle, J., Cress, U.: Group awareness and self-presentation in computer-supported information exchange. International Journal of Computer-Supported Collaborative Learning 3(1), 85–97 (2007)

8. Kolb, D.: Experiential Learning: Experience as the source of learning and development. Prentice Hall, New Jersey (1984)
9. McAuley, A., Stewart, B., Siemens, G., Cormier, D.: The MOOC Model for digital practice (November 2012)
10. Olofsson, A.D., Lindberg, J.O.: Whatever happened to the social dimension? aspects of learning in a distance-based teacher training programme. Education and Information Technologies 11(1), 7–20 (2006)
11. Rodriguez, O.: The concept of openness behind c and x-moocs (massive open online courses). Open Praxis 5(1) (2013)
12. Salmon, G.: E-tivities: The Key to Active Online Learning. E-tivities: The Key to Active Online Learning. Kogan Page, Limited (2002)
13. Siemens, G.: Connectivism: A Learning Theory for the Digital Age. International Journal of Instructional Technology & Distance Learning, 2 (2005)
14. Siemens, G.: Moocs are really a platform (2012), http://www.elearnspace.org/blog/2012/07/25/moocs-are-really-a-platform/ (accessed March 04, 2013)
15. Staemmler, D.: Individuelle Differenzen beim Lernen mit interaktiven Hypermediasystemen. Phd thesis, Universität Hamburg (2005)
16. Wenger, E.: Communities of Practice: Learning, Meaning and Identity. Cambridge University Press, Cambridge (1998)
17. Wheeler, S.: Recycling kolb (2012), http://steve-wheeler.blogspot.de/2012/06/recycling-kolb.html (last viewed June 18, 2013)
18. Willems, C., Meinel, C.: Online assessment for hands-on cybersecurity training in a virtual lab. In: Proceedings of the 3rd IEEE Global Engineering Education Conference (EDUCON 2012), Marrakesh, Morocco. IEEE Press (2012)

Supporting Citizen Inquiry: An Investigation of Moon Rock

Eloy David Villasclaras-Fernandez[1], Mike Sharples[1],
Simon Kelley[2], and Eileen Scanlon[1]

[1] Institute of Educational Technology, The Open University, UK
{eloy.villasclaras-fernandez,mike.sharples,e.scanlon}@open.ac.uk
[2] Faculty of Science, The Open University, UK
s.p.kelley@open.ac.uk

Abstract. Citizen inquiry is an innovative way for non-professionals to engage in practical scientific activities, in which they take the role of self-regulated scientists in informal learning contexts. This type of activity has similarities to inquiry-based learning and to citizen science, but also important differences. To understand the challenges of supporting citizen inquiry, a prototype system and activity has been developed: the *Moon Rock Explorer*. Based on the *nQuire* Toolkit, this offers people without geology expertise an open investigation into authentic specimens of Moon rock, using a *Virtual Microscope*. The *Moon Rock Explorer* inquiry has been evaluated in an informal learning context with PhD students from the Open University. Results of the evaluation raise issues related to motivation and interaction between inquiry participants. They also provide evidence that the integration of scientific tools was successful and that the *nQuire* Toolkit is suitable to deploy and enact citizen inquiries.

Keywords: Citizen inquiry, nQuire, Science learning, Informal learning.

1 Introduction

There is a growing interest in involving non-professionals and learners in scientific activities. For example, Alberts [1], in an editorial for the Science, argues that wider personal engagement in "carefully designed, hands-on, inquiry-based exploration of the world" (p. 1604) will inform public debate and may lead to scientific breakthroughs. A central challenge is how to enable widespread involvement in this scientific inquiry process, such that large scale online collaboration can be combined with inquiry based learning. We aim to address this challenge by developing **citizen inquiry** as a new theory and practice of informal education.

The aim of citizen inquiry is to create a novel synthesis of **citizen science** and **inquiry-based learning**. Since the early 20th century, citizen science projects have enabled non-professionals to participate in real scientific investigations, more recently through online activity. *Galaxy Zoo* [2] relies on volunteers doing analysis of astronomy images. In the *Great Sunflower Project* [http://www.greatsunflower.org/] members of the public can contribute to study bee populations by cultivating plants in their own gardens.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 383–395, 2013.

For *Fold it* people engage in creative game-like activities to propose new protein geometry [3]. The motivation for citizen science projects comes typically from the need for large amount of computer power (e.g., *Seti@Home*, *Rosetta@Home*) or intelligent effort, as in the aforementioned projects. As a by-product, citizen science allows people to become members of the scientific community and contribute towards the development of innovative science. It may also enable members of the public to learn about topics of their interest or gain understanding of scientific methods, but supporting productive learning is not integral to citizen science.

Inquiry-based learning (IBL) also has a long pedigree. From the early 20th century onwards [4] there have been proposals that children should learn science through collaborative inquiry. Since scientific thinking is essentially social, Dewey proposed that schools should become “laboratories of knowledge-making” [4, p. 127] where children engage in experimentation, communication, and self-criticism, constituting a youthful commonwealth of cooperative inquiry. Yorks & Kasl [5] have shown how scientific inquiry can be applied successfully in adult learning, as a systematic and productive process for learning from personal experience, consisting of repeated episodes of reflection and action through which a group of peers strives to answer a question of mutual importance. The inquiry-based learning process of proposing and conducting experiments, collecting data, and engaging in self-criticism based on one’s own data, can have positive effects on the understanding of content material and the scientific method [6].

To date, most inquiry-based learning activities have been conducted with pupils in the classroom, mediated by a teacher. Recent research has devised a new process of curriculum-based inquiry that extends beyond the classroom, supported by a personal inquiry toolkit implemented on a mobile device. In this approach, children typically start a science investigation in the classroom managed by a teacher, then continue it at home or outside with the aid of the mobile toolkit, then share, discuss, and present their findings back in the classroom [7]. This has shown success in engaging children with scientific activity, in teaching science topics, and in maintaining enjoyment of science lessons. But there is no evidence that children come to identify more strongly with science and scientists as a result of engaging with school-managed personal inquiry activities, nor of increased leisure interest or personal engagement with science outside school.

2 Citizen Inquiry

The term ‘citizen inquiry’ refers to the design and enactment of scientific projects by non-professional scientists in supportive communities, combining the benefits of large scale participation in authentic science practices with inquiry into scientific topics of personal interest and value. Our aim is to design and deploy an infrastructure that enables members of the public to develop their scientific interest in an innovative manner, in which they become independent scientists and investigate questions and hypotheses of their own or shared interest.

Citizen inquiry integrates methods from personal inquiry learning, citizen science, and open science [8]. It shares with citizen science an open nature, being driven by personal interest and developed outside of formal education. Compared to different models of citizen science [9], citizen inquiry is closer to participatory action research; however the goal is not to involve members of the public in professional scientific projects. Rather, it focuses on empowering members of the public to employ scientific tools and methods, augmenting their autonomy to plan and conduct scientific investigations not necessarily managed by professionals. It could be undertaken at home or outdoors, by children or adults, providing there is opportunity to enact a complete inquiry cycle, a rich physical or social environment to investigate, a community of engaged peers, and a set of tools to guide the process and collect data. These requirements present challenges of theory and practice:

- *Motivating participants*. Citizen science projects attract the interest of volunteers by their high profile, such as the search for extra-terrestrial life, or their contribution to science by addressing important medical or environmental problems. Conversely, students engaged in IBL activities are typically guided by their teachers, with extrinsic motivation provided by the curriculum and examining. We need to investigate whether an activity without these external influences will attract and maintain the interest of participants. In citizen inquiry, the initial motivation comes from personal interest in a topic, which is then maintained by forming or joining a club of investigators with similar interests and a diversity of contexts.
- *Devising scientific activities*. Devising a scientific question is a challenging task. While citizen inquiries may be developed out of curiosity without need for wider relevancy, we are interested in the possibilities for non-professionals to make valid contributions to science. This may not be the ‘big science’ of medical advances or scientific breakthroughs, but it should be personally relevant to the participants and also have a wider meaning and validity, adopting methods recognised by the scientific community. Investigations into supernatural, metaphysical or pseudo-scientific phenomena may not enable citizens to access scientific literature, learn how to use complex tools, understand scientific theories, or adhere to the ethical principles of the scientific community.
- *Managing inquiry processes*. Citizen science relies on professional scientists to plan investigations with suitable methodologies, whereas IBL activities are managed by some combination of teacher and students. Some give responsibility to the learners to plan their experiments; others encourage learners to decide the research questions that will be investigated. A goal of citizen inquiry is to offer learners the opportunity to propose and design their investigations. However, it is challenging to devise investigations that are personally engaging, testable by available tools, and sufficiently constrained to allow collection and analysis of data [7]. Typically, learners have problems managing their process of inquiry. They need specific support in designing appropriate experiments (e.g. what variables to choose, how to state and test hypotheses), implementing experiments (e.g. making predictions and avoiding being fixated with achieving particular results rather than testing hypotheses), and interpreting results (e.g. explaining graphs) [10-11]. The teacher has an essential role to provide guidance and to ensure appropriate quality of questions and methods planned by the students [12]. It is unclear whether non-scientists in *ad hoc* teams can plan or adopt inquiry processes that follow the good practices of professional science.

Within the IBL field, these issues are being tackled by the development of computer systems that provide support to learners, guiding them through a structured process of devising a topic, deciding a research question, planning a method, collecting and analysing data, answering the research question, sharing findings and reflecting on progress [13]. Following the same approach, we have developed a citizen inquiry support system, built on the *nQuire* inquiry learning platform, to manage the process of open participation in science inquiry and to study the role of technology in addressing the aforementioned challenges.

In addition to devising and testing software to support citizen inquiry, we are interested in understanding the management of online scientific communities. Online collaborative forums have been deployed to support inquiry learning [14] and are widely used in citizen science [15]. The *iSpot* platform has a reputation management system [16], enabling users to have their developing expertise recognized within the community. The *Fold it* user community provides a novel form of community engagement through protein folding competitions [3]. Given their importance in citizen science, we are interested in understanding how strong user communities can support motivation and learning in citizen inquiry if neither teachers nor science experts were available to guide the scientific process.

Our interest in citizen inquiry arises from its potential as a mechanism for informal learning, to facilitate comprehension of science in the everyday world, and the scientific methodology behind it. Specifically, the research presented in this paper is focused on how citizen inquiry activities can be supported through online systems and the building of online communities of users with shared interests. The next section describes the development the *nQuire* inquiry support system, already openly available, to support online investigation of lunar rocks: the *Moon Rock Explorer*. Section 4 describes the design of a study to evaluate this inquiry, which was conducted during February 2013. Section 5 discusses the results, and lastly, Section 6 presents the conclusions and future steps.

3 Development of a Citizen Inquiry Activity

Following the challenges listed above, this section describes a citizen inquiry on the subject of Moon Geology. Rather than developing a new software system, we have adapted *nQuire*, an existing tool for IBL.

3.1 Inquiry Based Learning Support

The *Personal Inquiry* project developed the *nQuire* toolkit to support IBL activities for science learning in secondary education [17]. The toolkit runs on laptop or tablet computers and enables teachers set up investigations that can be conducted within and outside the classroom. Investigations developed for the Personal Inquiry project covered the themes of Myself (activity and heart rate; healthy eating), My Environment (effect of noise pollution on bird feeding; microclimates; urban heat islands), and My Community (food packaging and decay).

A distinctive feature of *nQuire* compared to other IBL tools (such as *SCY* and *Let's Go*) is its representation of an entire investigation [18].

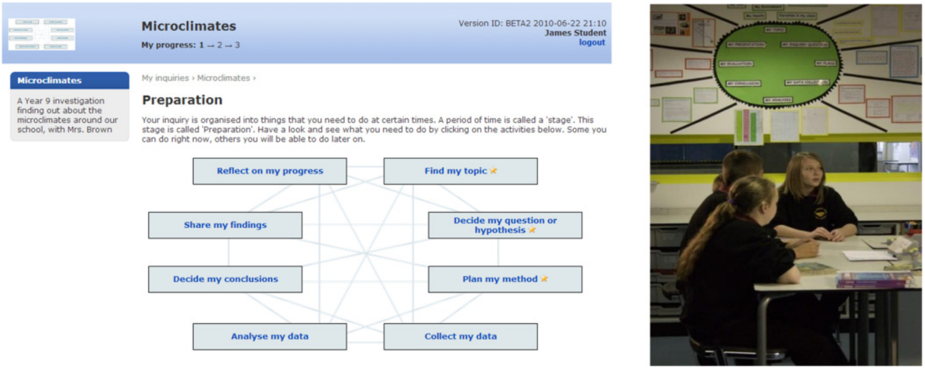


Fig. 1. Representation of the inquiry cycle (a) on the *nQuire* screen, (b) on the classroom wall

As shown in Figure 1(a), *nQuire* depicts an investigation as a cyclical sequence of activities. This is shown as an interactive diagram on the *nQuire* home screen and is also used by the teachers to organise the sequence of activities: one teacher produced a version of the diagram as a large poster on the classroom wall Figure 1(b). It shows the activities as interconnected, indicating that although students are expected to progress round the inquiry cycle, they may also start at any phase and revisit earlier phases, for example to revise the inquiry question so that it matches the methods.

Current development of *nQuire* is being continued through its integration into the *OpenScience Laboratory* (OSL) (http://www8.open.ac.uk/choose/ou/openscience), a project funded by The Open University and the Wolfson Foundation. The OSL is an online hub to access virtual instruments and practical science experiments. It integrates virtual and remote scientific tools, such as a field trip in a virtual world and control of remote telescope, accessible through web browsers. All the data provided by the real or virtual equipment is authentic, not simulated, gained from remotely-operated sensors, photo-realistic recordings of physical experiments and microscope images of real specimens. The OSL enables students on courses in higher education to conduct practical science experiments and some of the experiments will be open to the general public.

By integrating *nQuire*, the OSL will assist students and members of the public to engage in investigations that access sophisticated tools. This opens new possibilities for investigating IBL with *nQuire*, extending it from secondary education to undergraduate courses. It also represents an opportunity to investigate citizen inquiry.

3.2 The *Moon Rock Explorer* Citizen Inquiry

The *Moon Rock Explorer* is prototype demonstrator and test-site of citizen inquiry. It provides a self-managed investigation into Moon geology for people with no previous knowledge of Moon rock or geology. A user accesses the system through a public URL (*www.nquire.info/nquire*) and is presented with a video introduction to moon rock and an appeal to 'Investigate Moon Rock'. Having created an account, the user is taken to the home page, similar to that shown in Figure 1(a), but with six phases: Introduction, Decide my question, Plan my method, Collect my data, Analyse my data, Decide my conclusions. Entering at the Introduction phase, the user is shown photographs of four samples collected by Apollo astronauts – two of basalt Moon rock and two of Moon dust, or regolith – and is set the challenge to identify the differences between them.

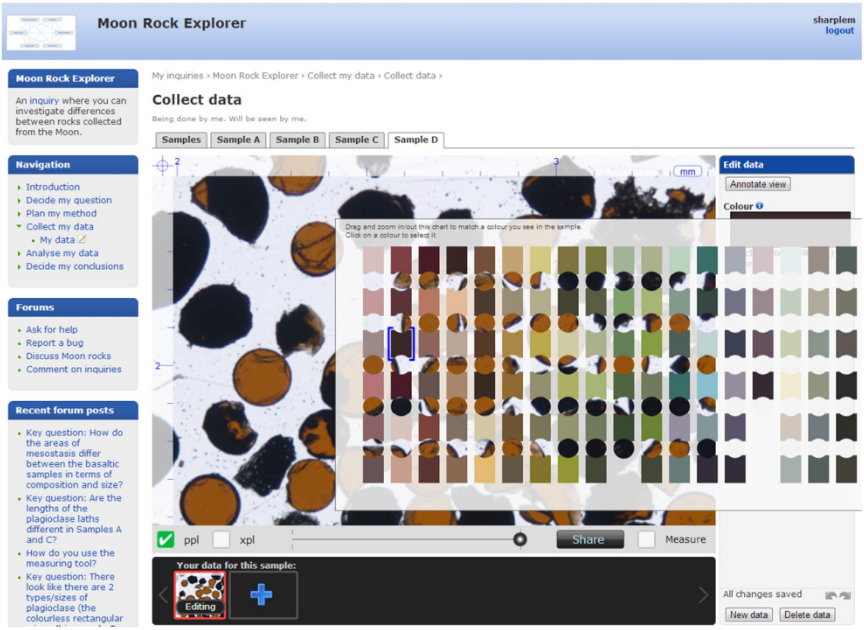


Fig. 2. Moon *Rock Explorer* inquiry, showing the *Virtual Microscope* tool and colour chart ('Drag and zoom this chart in and out to match a colour in the sample. Click on a colour to select it.').

3.3 User Experience of the *Moon Rock Explorer*

Typically, a user will first examine the four rock samples under a '*Virtual Microscope*' which allows study of digitised thin rock sections (0.03mm polished slices), with the ability to increase the magnification, zoom into an area of a slice and view in both plane-polarised and cross-polarised light. Then, the user will propose one or more specific questions to investigate, such as 'why are some grains in sample D bright orange?' or 'is there a difference in the average grain size between the four samples?' The questions need not be confined to the microscope, and the *Moon Rock Explorer* provides links to lunar maps and NASA sites giving background information on the Apollo missions. The next phase is to plan a method of investigation, which for the microscope involves selecting the measures from: grain size, distance, colour, opacity and crystal shape. The microscope provides tools to make these measurements for each sample. Having collected data, the user may move to analysing the data by creating a graph for selected measures (see Figure 3). Lastly, the user can propose answers to the questions based on a personal interpretation of the measures and plots, then post an automatically-compiled set of findings on the public forum.

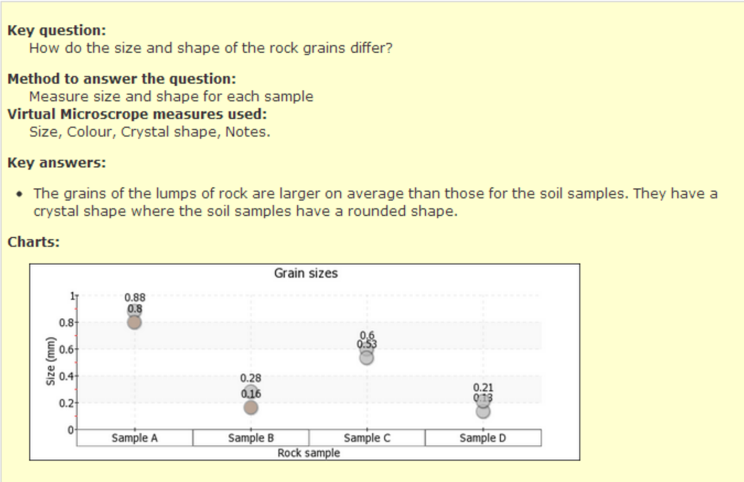


Fig. 3. Summary of an inquiry compiled by the *Moon Rock Explorer*

Although the *Moon Rock Explorer* is implemented on a version of the *nQuire* platform used for the Personal Inquiry project, it differs in three main aspects that are intended to support self-directed citizen inquiry rather than teacher-led inquiry:

1. **Simplified representation of the inquiry cycle**. The inquiry cycle was adapted for self-managed learning by adults, by reducing the number of phases (e.g., 'Reflect on my progress' was removed) and amending the informational text in each phase, accounting to take account of the fact that no teacher is present. Temporal restrictions on access to different phases of the inquiry were eliminated: the user can always access and re-visit any phase of the inquiry.
2. **Integration of scientific tools.** The *Moon Rock Explorer* was integrated with the Open University *Virtual Microscope* to promote scientifically relevant inquiries and learner motivation. Access to real Moon rock samples is extremely limited, so students and members of the public would not normally be able to examine them in detail. The microscope includes a measure tool, which lets users collect accurate data on grain size and separation, and a colour chart (Figure 2), based on a Rock-Color Chart produced by the Geological Society of America, to identify colours in the sample. Data collected with these tools is automatically recorded in *nQuire*.
3. **Addition of an online community system.** Although each investigation can be carried out autonomously, users can use the forum to interact with others. In addition, when an inquiry summary is generated by the system, there is a button to 'publish this question and discuss it with other users'. If a user chooses to do so, the system will create automatically a new forum thread to show with the investigation summary (see Figure 3). Changes to the research question, the research method, data, or answers, will also be updated in the forum thread. These threads can be used for discussion or even for peer-review between users.

3.4 Development of the *Moon Rock Explorer*

The *Moon Rock Explorer* was developed through a rapid iterative cycle of design, implementation and testing, initially involving members of the team, then with expert usability testers conducting heuristic evaluations [19]. The heuristic evaluation helped us to improve the user interface, especially to enable users to collect large amounts of data easily using the *Virtual Microscope*.

4 Evaluation

The development of the *Moon Rock Explorer* inquiry allows us to investigate the challenges of citizen inquiry activities. To that purpose, we have conducted a study using the *Moon Rock Explorer* inquiry in an informal learning setting. The goals of the study are to determine the relevance of the challenges discussed in Section 2 and assess the support implemented in the *Moon Rock Explorer* inquiry to overcome them:

- **Issue 1: Motivation**. Is the *Moon Rock Explorer* inquiry adequate for the Geology novices? Do the inquiry activities maintain participants' motivation? Do users support each other in their personal inquiries?
- **Issue 2: Management of inquiry processes**. Can participants develop complete inquiries following the inquiry cycle adapted from IBL activities? Is the guidance provided by the system sufficient?
- **Issue 3: Relevant scientific activities**. Are the questions proposed by the participants scientifically relevant? Does the integrated *Virtual Microscope* allow the participants to collect data for their investigations?

4.1 Method

The study is designed to simulate an informal learning activity involving PhD students in from the Faculty of Sciences of the Open University. To frame the study in an informal learning context, their participation was voluntary and the staff responsible for the study did not provide any guidance as in teacher-led activities. The study had the following structure:

1. They are invited to participate through email, including an activity schedule.
2. A one-hour face-to-face briefing session introduces them to the *Moon Rock Explorer*. Participants are encouraged to join, complete an inquiry, or just visit the forum.
3. The *Moon Rock Explorer* website is open for two weeks. Users can register and access any part of the system. Two academics from the Faculty of Science join the inquiry in the role of 'expert geologists', and monitor the forum for questions. Except for indicating that we will accept participation within a two-week period, there are no scheduled tasks, and the participants are free to register and complete the activities in their own time.
4. Online survey. A brief questionnaire about their perception of the experience.
5. Focus group. Participants are invited to a one-hour focus group, two weeks after the completion of Phase 3.

4.2 Sample

Participants were recruited from the collective of PhD students in the Faculty of Sciences at the Open University; their PhD studies are mostly related to astronomy, geology and planetary science. Therefore, the original participant pool included students who were interested in science and had differing levels of expertise in general geology and the geology of the Moon.

Six students eventually registered in the system and carried out (completely or partially) the inquiry. While the number of participants is lower than initially expected, it reflects the voluntary nature of the inquiry and the study.

4.3 Data Analysis

The goals of the study are focused on understanding the challenges of supporting this type of inquiry, rather than on observing gains on students' domain knowledge, scientific skills or attitudes towards science. The study was designed according to a mixed method that combines qualitative and quantitative data [20]. Both types of data are combined to reveal tendencies related to participants' activity and their interaction with the system and between themselves.

Qualitative data includes: participants' profile, answers to the online questionnaires (identified as *[OQ]* in the rest of the paper), answers in the focus group (*[FG]*), and contributions to the forum (*[F]*). Quantitative data is obtained from *nQuire* logs, which include the amount of user-created content and page visits.

5 Results

This section details the results obtained from the evaluation described in Section 4. The analysis of the data is presented here in relation to the issues discussed throughout the paper. Given the low number of participants, the objective of this section is to provide initial evidence of the relevance of the challenges discussed above, rather than to formulate conclusions.

5.1 Motivation

Motivation and entry barrier. The first question concerning motivation, as indicated in Section 4, was: "**Is the *Moon Rock Explorer* inquiry adequate for the Geology novices?**" Only one student without previous experience in geology joined the inquiry, and this participant created a significantly lower amount of inquiry content (see Table 1). Also, against our expectation, the 'expert participants' were more interested in the inquiry than providing help to other users (see 'motivation and user interaction' below). These results suggest that the *Moon Rock Explorer* has a high entry barrier for the target audience, and was more suitable for experienced geologists.

Motivation and Inquiry Design. To answer the question **"Do the inquiry activities maintain participants' motivation?**" we consider two different issues.

Table 1. Amount of content created by each participant. Experience: participants prior knowledge (*E*: doing research on geology; *L*: different background, but limited experience in geology; *N*: no previous experience). Key questions, Data, Data charts and Key answers: number of items of created by each participant.

Participant	Experience	Key questions	Data	Data charts	Key answers
1	*E*	1	35	1	0
2	*E*	1	32	1	0
3	*E*	1	0 (used external software)		1
4	*E*	2 (1 investigated)	15	0	0
5	*L*	3 (1 investigated)	40	3	1
6	*N*	1	0	0	0

First, the Moon Rock Inquiry was designed as an open-ended inquiry without fixed research questions. While this may have contributed to the high entry barrier, it was suitable for participants with prior knowledge of geology. This is supported by the comments in the final questionnaire (e.g.: "*I liked being able to explore the four samples with the virtual microscope […]*" [OQ]), the quality of the questions they proposed, and by the amount of data they collected to answer their questions (see Table 1), amounting to two hours work in the case of one participant [FG].

A second issue goes beyond the selected domain of inquiry. Some participants did not answer any questions, even after collecting data for them (Table 1). Asked about this, the participants found that there was a lack of purpose in the inquiry [FG]. Indeed, the only product of their activity that could be shared with other users was the automatic forum investigation summaries, which had problems discussed below.

Motivation and User Interaction. The objective of investigating interactions between users was obviously affected by the low number of participants. Users could interact through the forum, even though the inquiry activities were individual. We have identified and characterized a number of interaction types.

1. The first type of interaction is related to participants **discussing the investigations of others in the forum**. Unfortunately, no participants commented on other investigations, which was an unexpected result. They reported that they did not feel comfortable to do so, because they did not know the state of each investigation, and were concerned about intruding [FG].
2. Participants did, however, **visit other investigations**. This happened 31 times. Log analysis allows us to identify two patterns in this behaviour, as shown in Figure 4. The first pattern (Figure 4-left) was carried out by three participants with previous expertise in Geology. They did not look at other investigations, even though they visited forum topics related to bugs and the tool. The second pattern (Figure 4-right) was followed by three participants with varying levels of expertise. Again, this was against our expectation, as we had expected interest from the experts to see other participants' work and help, if possible.
3. Related to this, a third type of interaction involves **asking domain related questions** and obtaining answers from knowledgeable members of the community. This happened only once [F]. The question was answered by a member of staff, and the rest of the participants did not intervene in the discussion.

4. A final form of interaction consists of **asking non-domain questions**. This was limited to bug reports made in the forum, which were answered by the administrator after the problems had been fixed [F].

Unfortunately, the results indicate that the system did not encourage participants to interact with each other. The available data indicates thus a negative answer to the question: "**Do users support each other in their personal inquiries?**"

5.2 Management of Inquiry Processes

The management of inquiry processes had been evaluated thoroughly using *nQuire* in teacher-led inquiries. In contrast, in our study the role of a teacher or tutor was not present, so we were concerned about a potential lack of support.

The analysis of the participants' work indicates however that the participants could complete the whole activity, leading to positive comments such as: "*I also liked the way that each part of the investigation led on to the next,*" [OQ]. Thus, we answer positively to the questions "**Can participants develop complete inquiries following the inquiry cycle adapted from IBL activities?**" And: "**Is the guidance provided by the system sufficient?**"

Besides the support provided to guide participants across the proposed inquiry process, we are interested in their own processes. The *nQuire* Toolkit was designed to allow participants to follow different paths, visiting the activities in their order (the Network representation in Figure 1). However, the log analysis shows that nearly all the participants progressed in the order suggested by the Navigation element of the *nQuire* user interface (left panel in Figure 2). The only case of an alternative process is shown in Figure 5. This participant started by collecting data, in order to familiarize herself with the tool, then moved to propose a key question. Participants noted contrasts between the Network and linear Navigation representations and were confused by the differences [FG]. It is important to understand whether the *nQuire* interface discourages alternative progressions, and the potential risks of this issue.

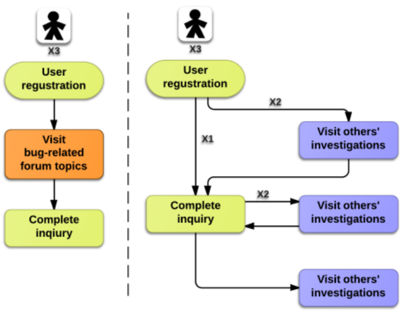


Fig. 4. Patterns of interaction through investigation visits. Number of participants is indicated

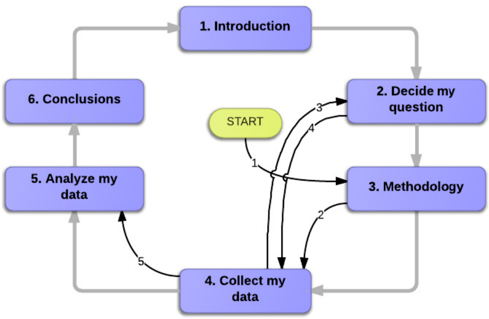


Fig. 5. Alternative participant learning flow. Grey lines represent the suggested path; black lines the alternative path.

5.3 Relevant Scientific Activities

Regarding the last issue, we were interested in whether "**the integrated *Virtual Microscope* allow the participants collect data or their investigations**" The amount of data collected by the participants (Table 1) confirms that they were capable of collecting a large number of measures using the *Virtual Microscope*. Qualitative data confirms this, e.g.: "*comparing the four slides quickly and easily - more easily, in fact, than would be possible with a real microscope!*" [OQ]

The study revealed, on the other hand, technical problems. First, the participants reported that after a number of data had been collected, it become difficult to remember which features of a specimen had been measured [FG][OQ]. A second problem is the lack of a powerful data analysis tool: one participant collected her data using external spreadsheet software that allowed more sophisticated analysis [F]. Despite these problems, we consider that the evaluation of the *nQuire/Virtual Microscope* integration is positive, as it satisfied its goal: to provide a tool to collect scientifically reliable data which can be used to answer research investigations.

6 Conclusions and Future Directions

The development and evaluation of the *Moon Rock Explorer* inquiry is the first step towards understanding the requirements of citizen inquiry. The long term goals of this research are to support learners in adopting an inquiry-based learning approach in a self-regulated and self-directed context, and to support the creation of citizen inquiry user communities. A milestone in this research is the integration of *nQuire* with the scientific instruments of the OSL, which will enable users to create inquiries for a wide range of scientific domains.

The study presented in this paper provides initial evidence of the suitability of *nQuire* and the *Virtual Microscope* to support citizen inquiry activities, notwithstanding some problems with the current implementation that need to be addressed. A key result of the study is the successful guidance provided by *nQuire*. The tool is currently being extended with an inquiry authoring tool. Feedback from the study is being used to refine the *nQuire-Virtual Microscope* user interface.

The most exciting open challenges are related to the issue of motivation. Several mechanisms are under consideration to foster motivation in collaborative inquiries: reputation systems (as in *iSpot*), support for roles within large scale inquiries (similar to citizen science projects), and an internal peer-reviewed 'scientific journal' for members of the citizen inquiry community. Further research is needed to understand these and other mechanisms in the context of citizen inquiry.

Acknowledgements. This work is developed within The OpenScience Laboratory project (www.open.ac.uk/blogs/openscience/), funded by The Open University and the Wolfson Foundation. The authors would like to thank the OU Faculty of Science academics and students that participated in the study; members of the Institute of Learning Technology that participated in the evaluation of the Moon Rock Explorer inquiry; and members of the Personal Inquiry project (www.pi-project.ac.uk/) from University of Nottingham and the Open University for their work the nQuire.

References

1. Alberts, B.: Editorial: Science Breakthroughs. Science 334(6063), 1604 (2011)
2. Raddick, J., Lintott, C.J., Schawinski, K., Thomas, D., Nichol, R.C., Andreescu, D., Bamford, S., et al.: Galaxy Zoo Team: Galaxy Zoo: An Experiment in Public Science Participation. Bulletin of the American Astronomical Society 38, 892 (2007)
3. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., et al.: Predicting protein structures with a multiplayer online game. Nature 466(7307), 756–760 (2010)
4. Dewey, J.: Science as Subject-matter and as Method. Science 31(787), 121–127 (1910)
5. Kasl, E., Yorks, L.: Collaborative inquiry for adult learning. New Directions for Adult and Continuing Education 2002(94), 3–12 (2002)
6. White, B.Y., Frederiksen, J.R.: Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. Cognition and Instruction 16(1), 3–118 (1998)
7. Anastopoulou, S., Sharples, M., Ainsworth, S., Crook, C., O'Malley, C., Wright, M.: Creating personal meaning through technology-supported science inquiry learning across formal and informal settings. International Journal of Science Education 34(2), 251–273 (2012)
8. Woelfle, M., Olliaro, P., Todd, M.H.: Open science is a research accelerator. Nature Chemistry 3, 745–748 (2011)
9. Wilderman, C.C.: Models of community science: design lessons from the field. In: McEver, C., Bonney, R., Dickinson, J., Kelling, S., Rosenberg, K., Shirk, J.L. (eds.) Citizen Science Toolkit Conference. Cornell Laboratory of Ornithology, Ithaca (2007)
10. van Joolingen, W., de Jong, T., Lazonder, A., Savelsbergh, E., Manlove, S.: Co-Lab: research and development of an online learning environment for collaborative scientific discovery learning. Computers in Human Behavior 21(4), 671–688 (2005)
11. Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R.G., Kyza, E., et al.: A Scaffolding Design Framework for Software to Support Science Inquiry. The Journal of the Learning Sciences 13(3), 337–386 (2004)
12. Sharples, M., Anastopoulou, S.: Designing orchestration for inquiry learning. In: Littleton, K., Scanlon, E., Sharples, M. (eds.) Orchestrating Inquiry Learning, pp. 69–85. Routledge (2011)
13. Linn, M.C., Clark, D., Slotta, J.D.: WISE Design for Knowledge Integration. Science Education 87(4), 517–538 (2002)
14. Linn, M.C., Slotta, J.D.: Enabling participants in online forums to learn from each other. Collaborative Learning, Reasoning, and Technology, 61–97 (2006)
15. Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., Shirk, J.: Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. BioScience 59(11), 977–984 (2009)
16. Clow, D., Makriyannis, E.: iSpot Analysed: Participatory Learning and Reputation. Learning, 34–43 (2011)
17. Mulholland, P., Anastopoulou, S., Collins, T., Feisst, M., Gaved, M., Kerawalla, L., Paxton, M., et al.: nQuire: technological support for personal inquiry learning. IEEE Transactions on Learning Technologies First publ.(99), 1–14 (2012)
18. Scanlon, E., Anastopoulou, S., Kerawalla, L.: Inquiry learning reconsidered: contexts, representations and challenges. In: Littleton, K., Scanlon, E., Sharples, M. (eds.) Orchestrating Inquiry Learning, pp. 69–85. Routledge (2011)
19. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proc. ACM CHI 1990 Conf., Seattle, WA, April 1-5, pp. 249–256 (1990)
20. Cairns, P., Cox, A.L.: Research methods for human-computer interaction. Cambridge University Press (2008)

Affective Metacognitive Scaffolding and Enriched User Modelling for Experiential Training Simulators: A Follow-Up Study

Gudrun Wesiak[1,2], Adam Moore[3], Christina M. Steiner[1], Claudia Hauff[4], Conor Gaffney[5], Declan Dagger[5], Dietrich Albert[1,2], Fionn Kelly[6], Gary Donohoe[6], Gordon Power[5], and Owen Conlan[3]

[1] Knowledge Technologies Institute, Graz University of Technology, Austria
[2]Department of Psychology, University of Graz, Austria
{gudrun.wesiak,christina.steiner,dietrich.albert}@tugraz.at
[3] KDEG, School of Computer Science and Statistics, Trinity College, Dublin, Ireland
{mooread,owen.conlan}@scss.tcd.ie
[4]Delft University of Technology, the Netherlands
c.hauff@tudelft.nl
[5] EmpowerTheUser, Trinity Technology & Enterprise Campus, The Tower, Dublin, Ireland
{conor.gaffney,declan.dagger,gordon.power}@empowertheuser.com
[6]Department of Psychiatry, School of Medicine, Trinity College, Dublin, Ireland
fkelly@stpatsmail.com, donoghug@tcd.ie

Abstract. The ImREAL project is researching how to meaningfully augment and extend existing experiential training simulators. The services developed support self-regulated, goal-, and application-oriented learning in adult training. We present results from a study evaluating a medical interview training simulator that has been augmented by an affective metacognitive scaffolding service and by user modelling exploiting social digital traces. Data from 152 medical students participating in this user trial were compared to the results of a prior trial on an earlier technology version. Findings show that students perceived the learning simulator positively and that the enhanced simulator led to increased feelings of success, less frustration, higher technical flow, and more reflection on learning. Interestingly, this cohort of users proved reluctant to provide their open social IDs to enrich their user models.

Keywords: training simulator, self-regulated learning, affective metacognitive scaffolding, user modelling, social digital traces, evaluation.

1 Introduction

Today's technology-enhanced learning market offers a variety of adaptive cognitive systems that automatically recognize individual learners' needs and are able to create engaging and motivating learning experiences. In this context, experiential training simulators are gaining increasing popularity and importance as learning technologies for adult training. Adult learning is characteristically self-directed, experienced-based,

D. Hernández -Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 396–409, 2013.

goal-oriented, intrinsically motivating, and relevant to the real world application context [1]. Immersive simulated environments for experiential learning need to correspond to these principles of andragogy by creating a situational context in which diverse kinds of skills relevant for job practice can be acquired and practiced [2]. Additionally, training environments should be tailored to the individual, taking into account each learner's goals, prior experiences, and metacognitive competence [3]. This makes self-regulation and metacognition, which have a long research tradition in pedagogy and psychology (e.g. [4],[5]), key topics in the definition of sound conceptual underpinnings of learning technologies, in general, and training simulators, in particular [3],[6].

The EU-funded ImREAL project[1] is researching how to augment existing simulators to increase their meaningfulness and their relevance by real-world activity modelling, enriched user modelling, and affective metacognitive scaffolding.

This paper is a follow-up to earlier work on developing, integrating, and testing a metacognitive scaffolding service (MSS) in a medical interview training simulator [7],[8]. After the initial evaluation of the training simulator (section 2) with integrated ImREAL services showed positive effects on learning motivation and perceived performance, the simulator augmentation has been further expanded: the metacognitive scaffolding service has been refined and also extended by incorporating affective aspects and the user modelling has been enriched by exploiting social digital traces as a basis for adaptation (section 2.1). Herein, we present the related ImREAL services and how their integration in the medical training simulation is shown to add value to simulation development and learning experiences. A user trial has been conducted to investigate the effects of the updated and extended augmentation. The results are compared to the previous user trial with metacognitive scaffolding only. Potential improvements to the prior version regarding self-regulated learning support, learning experience, learning performance, motivation, and service integration are analysed (section 3). Obtained outcomes and their implications for future work are discussed in section 4.

2 Experiential Training Simulator and Augmentation

EmpowerTheUser (ETU), a Dublin-based SME, has developed the innovative RolePlay Simulation Platform, empowering training experts with easy to use, rapid tools for developing immersive simulations. The platform provides a test-bed for the simulations in the ImREAL project and consists of three core tools: a Simulation Development Tool, an adaptive RolePlay Simulator, and an Analytics Dashboard. The Platform is actively used in training areas such as sales, customer services training, leadership, management and clinical interviewing (Fig. 1). Each RolePlay Simulation supports two running modes: assess mode (where all learner decisions are scored and a detailed performance report is presented) and practice mode (where learners can explore the scenario with all the learning interventions like coaching activated). With its assess mode the platform also serves as a psychometric profiling, behavioural measurement and skills assessment tool.

[1] `http://www.imreal-project.eu`

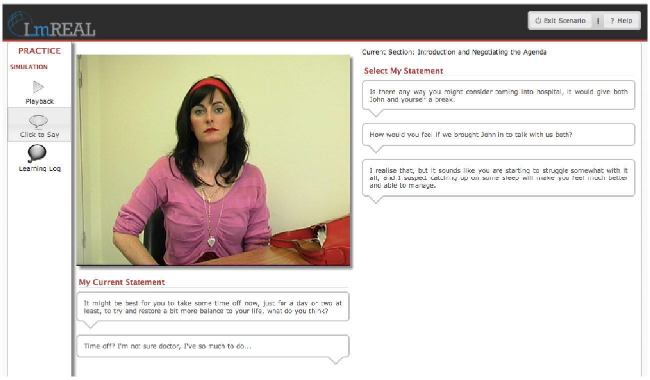


Fig. 1. ETU medical simulations: Interviewing a manic bipolar patient

2.1 Augmenting the ETU Simulator

During the ImREAL project a subset of the platform's learning triggers were extended to facilitate augmentations by the ImREAL suite of learning services, namely: coaching trigger (provides in-context feedback on the decision path) and reflection trigger (supports critical thinking and the transfer of learning). These learning triggers are modelled and configured *before* running a simulation by the subject expert/simulation developer. The focus in this paper is the impact of the ImREAL augmentations on ETU's RolePlay Simulation Platform in the context of medical training, particularly teaching student doctors about effective doctor-patient communication. ETU works closely with psychiatry and psychology experts at Trinity College Dublin's Medical School to create simulation scenarios for doctor-patient communication ranging from medical health to general medicine. Two enhancement opportunities are possible with the ImREAL suite of learning services: 1) augmenting reflection triggers with dynamic metacognitive scaffolding and 2) augmenting coaching triggers with more culturally oriented content.

To investigate the educational relevance of the simulator after implementation of these services, the following evaluation questions were addressed:

- Referring to the requirement of learning simulations providing authentic situational contexts [2],[3]: Does the simulation provide learning experiences that are *relevant* for users?
- Relating to the simulator augmentation: Are the *services* well *integrated* in the simulator and learning experience?

From Metacognitive Scaffolding to Affective Metacognitive Scaffolding. It has previously been described [8] how scaffolding is an important part of the educational process, supporting learners in knowledge acquisition and learning skill development.

As with the previously reported trial, metacognitive scaffolding was provided using calls to a RESTful service developed as part of the ImREAL project. The service utilises technology initially developed for the ETTHOS model [9] and presents items from the Metacognitive Awareness Inventory [10]. In an evolution from last year's scaffolding, however, prompts are displayed according to an underlying map created by the simulation's instructional designers.

It is commonly acknowledged that emotions play an important role in learning, exerting effects on information processing and performance [11]. In order to improve the provision of scaffolding to the learner, the existing MSS has been extended by an additional, affective element, thus becoming the Affective Metacognitive Scaffolding Service (AMSS). To obtain information of users' current affective state, an explicit Smiley-Based Affect Indicator (SBAI) [12] was inserted into the simulation feedback interface (*cf.* Fig. 2) and the enriched analysis and user modelling services, as described in the next section, were also utilised to provide additional information, where available, about the learner and their input.

With affect now included in the user model, and with a richer understanding of the learners, also the selection of the most salient metacognitive prompts to be displayed was changed. More priority was given to the weightings provided by the instructional designers. If high confidence values for learner affect attitudes were available, these were used to prioritise appropriate prompts (e.g. a prompt considered as encouraging might have a more prominent weighting for learners with negative affect indication).

The augmentation services (as presented below) are the basis for providing additional information to the learner model. If a learner provides a Twitter ID and textual input into the system, an affect stereotype is derived from sentiment analysis of their tweet stream or entry to update the model in real time. Affect stereotypes are positive (i.e. optimistic), neutral or negative (i.e. pessimistic). The augmentation service also provides a unique ID for each learner, which allows AMSS to send back the user model values that it derived and updated.

In addition, the presentation of the scaffolding service was slightly changed in this version of the simulation, with the open text box for collecting reflection consistently prefaced with a short text: "Reflect now on your learning: Was this last part of the simulation useful for you?", as shown in Fig. 2. Also, as the scaffolding was now provided as an XML bundle, styling was delegated to the simulation presentation engine, resulting in a more completely integrated interface and styling.

To integrate AMSS, ETU's Platform was extended in a number of ways: (a) The Simulation Development Tool was extended to accommodate mapping between simulation sections and the thinking prompts of AMSS. It also allows a reflection trigger to be tagged as dynamically "metacognitive". (b) The adaptive RolePlay Simulator can then interpret these new metacognitive reflection triggers, access the AMSS service and present the metacognitive prompts *in-situ*. (c) As part of the simulation self report dashboard, learners also have access to the AMSS SBAI affect reporting tool.

In the evaluation of AMSS the following questions were addressed in our study:

- Is *self-regulated learning* supported?
- Does the AMSS lead to better *learning experience?*
- Does the AMSS lead to a better *learning performance?*
- Does the AMSS enhance *motivation/affect*?

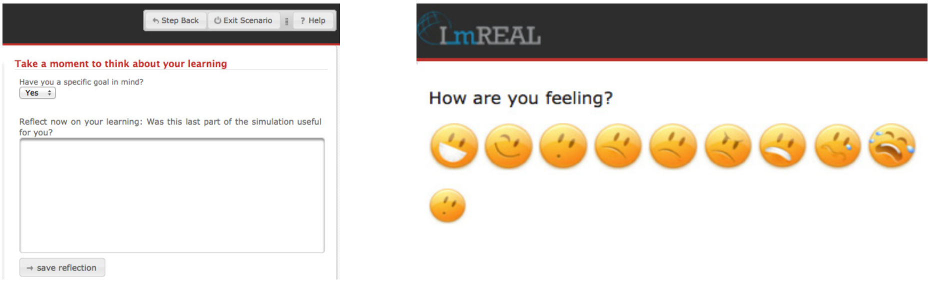


Fig. 2. Reflection trigger augmentation: Metacognitive scaffolding and capturing affective state

Enriched User Models Based on Digital Traces. The adaptation and personalization of e-learning systems requires user modelling, which involves collecting user information either explicitly, by asking the users a series of questions about their knowledge levels and skills or implicitly, by deriving a user profile from already existing data. Regarding the latter, digital traces on social Web platforms may give insights about relevant learner characteristics. With Facebook counting more than a billion monthly active users, Twitter and LinkedIn over 200 million users each[2], it is likely that today, most users of an e-learning system are regularly active on one (or several) social Web platforms. Exploiting the increasing amount of these digital traces learners create and share on social media is a potential resource for addressing the fact that adult learners arrive at a learning situation with highly diverse motivations and aims, real-world experiences and interests, knowledge and competence levels. Since digital traces are publicly available on the Web, we hypothesised that there will be few trust and privacy concerns for the users.

In our work on user model augmentation, we have investigated the utility of extracting different types of information nuggets about a learner from his social Web profiles, such as for example, the learner's geographic history (where has the learner been) [13], the learner's academic knowledge (which topics does the learner know something about) [14], the learner's high-level interests [15], and sentiment expressed in social media [16].

The U-Sem framework [17] developed in the ImREAL project is a service infrastructure for enriching and mining user data. Work on U-Sem focuses on how to derive user profiles from the social Web, how to enrich existing user profiles with that information, and how to verify the correctness of certain aspects of the automatically derived user models. U-Sem allows researchers and developers to create and design Web services for enriching and analysing user-related data, which are then made available to client applications, for instance a location detection service [13] used as a basis for providing adaptive, culturally-oriented coaching triggers.

To integrate U-Sem into the ETU simulator, its platform was extended in a number of ways: (1) The Simulation Development Tool was extended to support a

[2] See `http://newsroom.fb.com/Key-Facts`, `http://blog.twitter.com/2013/03/celebrating-twitter7.html`, and `http://press.linkedin.com/about` for Facebook, Twitter, and LinkedIn.

range of adaptive dimensions supported by the ImREAL services (e.g. cultural awareness). It also allows a coaching trigger to be tagged based on a specified adaptive dimension; (2) The adaptive RolePlay Simulator provides a mechanism for capturing a user's social tags (e.g. Twitter name or Flicker ID), registering them with ImREAL services and then adapting the coaching triggers to the user.

To evaluate this conceptual and technical approach for enriching user models the following questions were of interest:

- Do learners actively *use* one or more *social network(s)*?
- Do learners *feel comfortable* to freely provide their open social IDs?
- Does the augmented user model *correctly reflect* user characteristics?

3 Empirical Study

An empirical study was conducted with the goal of investigating the above outlined evaluation questions on the quality and impact of the service augmentation to the simulator. Overall 152 students from Trinity College, Dublin took part in this second user trial (UT2) and used the medical interview training on the ETU simulator with integrated AMSS and U-Sem services. Data collection was done during (log data) and after (questionnaire data) the training with the simulator. In addition, a cohort characterisation survey (on demography, metacognition, social network use, etc.) was conducted ahead of the trial.

3.1 Method

Instruments. Aligned with the evaluation questions outlined in Section 2 the following instruments were used:

Real-world relevance of the simulation was evaluated by two survey items.

Service integration of the ImREAL services in the simulator was measured by specific questions regarding the integration of AMSS and the flow of the simulation and learning experience (from a technical perspective).

Self-regulated learning (SRL) was assessed via an SRL questionnaire [18] with nine subscales measuring the general use of cognitive, metacognitive, and resource management strategies. In addition, interaction data and text entries tracked by the simulator served to investigate SRL behaviour during the training.

Learning experience covered workload and flow experience as aspects referring to how users perceive and experience the training with the simulator. These were measured with the NASA-TLX [19] and the Flow Short Scale [20], respectively. In addition, it was assessed how users experienced the thinking prompts provided by the AMSS, via a set of items provided in the post-simulation survey.

Learning performance in the simulator was evaluated via the objective assessment procedures built into the ETU simulator. Thereby, dialogue scores relating to the individual steps of the medical interview scenario and competence scores on several interview skills were used. The scoring constructs are derived from the skills defined in the Calgary-Cambridge model for medical interviewing. The constructs are then loaded in the simulation model and each decision path loads a particular weighting onto the respective skill constructs that are represented by a particular decision. All weightings across all paths taken by the learner during an attempt are then computed to produce the learners' performance score. The subscale 'performance' of the NASA-TLX served as additional subjective measure.

Motivation and affect. State motivation was explicitly queried by four specific questions; for affect assessment the Total Affective State Scales [21] were used as a subjective self-report. In addition, the SBAI of the AMSS provided the opportunity of capturing affect self-reports during the learning activity.

Questions on *social network use* and attitudes relating to *user privacy and trust* in the context of providing open social IDs for personalisation of the learning experience were part of the cohort characterisation study. Social IDs were queried in the cohort survey, as well as directly in the simulator for user model enrichment.

Participants and Procedure. UT2 participants were on average 22.81 years old (*SD* = 3.79) ranging from 19 to 45 years (note: *N*=95 for the cohort survey) and with equal gender ratio. Students participated from February until March 2013 as part of their curriculum. Before starting their training with the simulator participants completed the cohort characterization survey. During the training students could use the simulator as long and as often as they wished and deal with two psychiatric scenarios: mania and depression. Both scenarios could be entered in assess as well as practice mode. ImREAL services are integrated in the latter, whereas learning performance scores were assigned in assess mode. When entering the assess mode students were asked to predict their own scores regarding their performance in different steps of the clinical interview (introduction, eliciting information, outlining a management plan, closing the interview) and for interviewing skills (e.g. empathy, communication). The average overall duration of interacting with the simulation was 27.24 minutes (*SD*=11.03). After finishing their training, students provided their feedback on the simulator via an online survey covering the instruments listed above. Completion time of this survey was on average 10.75 minutes (*SD*=4.59).

3.2 Results

From the 152 participants using the simulator in the assess mode, different subsamples answered the cohort characterization and the post-simulation survey. Thus, sample sizes vary throughout this section. Where applicable results of this user trial (UT2) are compared to those of the prior user trial (UT1; 143 students using the simulator with MSS only in 2011/2012) reported in [8]. Comparisons are therefore based on independent samples, however participants in both studies were medical students from Trinity College, Dublin, with no significant differences regarding their sex or age. Log data was collected separately for the two scenarios. A scenario-specific analysis had shown that the scenarios were used to a different extent in the two user trials, but closer analysis yielded the same trends with respect to data distribution across the interview phases and the prediction scores and ETU scores per trial. Therefore, the data was aggregated across the two scenarios.

Real World Relevance. In both trials, participants were asked to (1) rate how relevant they think the experience with the ETU simulator was as preparation for real clinical interviews and (2) how well prepared (in percent 0-100) they felt for the interview based on their experience with the ETU learning environment. With a median of 3 on a 4pt. scale, the 40 participants from UT2 viewed the experience as rather relevant for their future interviews with patients, which does not differ significantly from UT1 ratings. With an average score of 59 (SD=16.16; N=37) on the second question UT2 students evaluated the relevance of the simulator higher than those of UT1 (M=51.3, SD=17.8; N=33), however the difference did not reach statistical significance ($T_{(68)}$=1.9, p=.062).

Service Integration. Integration of the metacognitive scaffolding, i.e. the thinking prompts is seen as neutral to positive. The perception of technical flow increased from UT1 to UT2 (M_{UT1}=3.75, SD=1.16; M_{UT2}=4.69, SD=1.15; $T_{(75)}$=-3.551, p=.001) indicating an appropriate integration of the service in the simulator.

Self-Regulated Learning. Data (N=25) from the nine SRL subscales range between means of 56.42 (SD=11.9) for 'memorising strategies' and 71.74 (SD=10.1) for 'elaboration strategies'. A MANOVA for independent samples revealed no difference between the two user trials, neither for the nine univariate nor for the multivariate analyses (for the latter $F_{(9,36)}$=.623, p=.77). Since for both trials all subscale means are above 50%, students from all samples show an average to good use of cognitive and metacognitive strategies (higher values indicate better strategies).

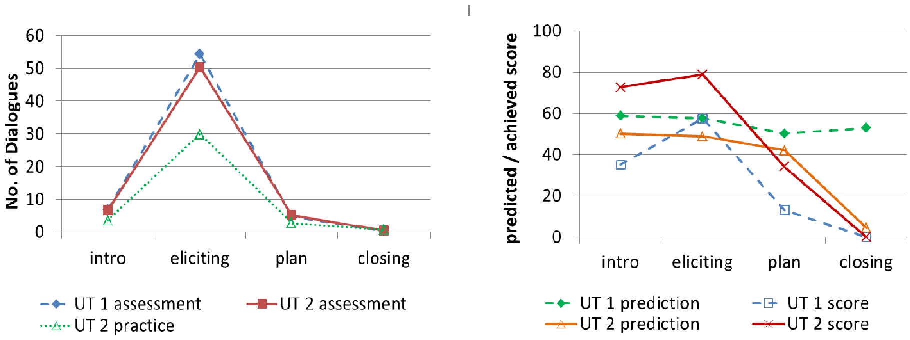


Fig. 3. Number of dialogues (left), predicted and achieved scores (right) in User Trials 1 and 2

Data on effort in terms of time spent with the simulation and number of dialogues were logged by the simulator. Because both measures are correlated (r=.487 for assess and .8 for practice mode, p<.001), and only the number of dialogues is reported here. Fig. 3 (left hand) compares the number of dialogues in different steps of the clinical interview during the assess modes of UT1 and UT2 and the practice mode of UT2. Mann-Whitney U-tests revealed no difference between the two user trials, irrespective of the step within the interview. For UT2, the assess mode was used more often and also more intensively than the practice mode. From the 152 students using the assess mode, 71 entered the practice mode and 58 processed dialogues as well. On the right hand side of Fig. 3, scores predicted by the students before entering the interviews and those actually obtained during the interview are shown. Whereas students from UT1 overestimated their performance for the introduction and planning phases, UT2 students underestimated their performance for the introduction and eliciting phases and overestimated it for the planning phase (*T*-tests for UT1: all $T_{(17)} \geq 6.5$, $p<.001$; Wilcoxon-tests for UT2: all $z \geq 3.2$, $p \leq .001$).

Another indicator for SRL strategies is the kind of information students recorded during the learning process. Overall 1092 thinking prompts were provided during the second user trial (M=15 per student, SD=16.9). With each prompt a text entry field popped up for collecting users' reflections (AMSS text entries) and students were asked about the usefulness of the last part of the simulation. From 69 yes/no responses 88.4% were positive. Additionally, the ETU simulator's notepad could be used at any time to reflect or take notes. Table 1 compares the number and types of UT2 notes (taken via AMSS or ETU) to those of UT1, coded by one researcher assessing the content of each entry and assigning its content to at least one of four types – Position, Technical, Patient Notes and Reflection [22]. Where the content was of more than one type, each was coded against that entry. The proportion of entries that were actually reflective increased in UT2, while the portion of all other entry types (positional, technical, and patient notes) decreased. Thus the type of notes taken by the students supports the assumption that AMSS fosters metacognition and reflection.

Table 1. Rounded percentages of content types for entries from the note-taking tools.

	Users	Text entries	Position	Technical	Notes	Reflection
UT 1	50	107	17	57	16	66
UT 2	35	86	1	13	7	93

Note: Entries can be coded to more than one type, thus percentages may exceed 100.

Learning Experience. 21 students of UT2 answered 10 questions on the perception of how helpful the *thinking prompts* were. On 5-pt. rating scales (from '1 not at all' to '5 very much') the overall score reached 3.55 (SD=.72), single scores ranged between 3.10 and 3.81. This indicates that the prompts are helpful with regard to content, timing, support in planning, monitoring, improving, and analysing one's learning performance. Comparisons with UT1 showed no significant differences between the overall or single item scores (all p >.2 for unrelated samples *T*-tests).

Two further measures used for learning experience are workload and flow. Average NASA-TX *workload* scores for the two user trials are depicted in Fig. 4. Multivariate MANOVA results revealed significantly lower overall workload for UT2 (multivariate $F_{(6,52)}$=7.6, p<.001). Results from single subscales indicate that the training in the simulator leads especially to load for effort and mental demand. Univariate analyses yield effects on performance and frustration ($F_{(1,52)}$=48.4, p<.001 for performance and $F_{(1,52)}=7.3$, p=.009 for frustration), indicating that students in UT2 felt less frustrated and had a stronger feeling to be successful than UT1 students.

Ratings for flow (on 7-pt. scales, with higher ratings indicating higher fluency and smoothness of the learning process and higher involvement in the task) from 40 UT2 users were compared to UT1 (N=37). The average UT2 rating of 4.75 (SD=.79) for overall flow was significantly higher than in UT1 with M=4.35 (SD=.91; $T_{(75)}$=-2.12, p=.038).

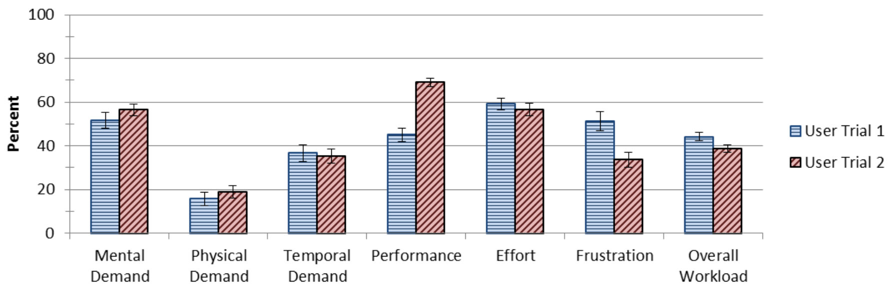


Fig. 4. Mean scores (and SD) from the NASA-TX workload scales for the two user trials

Learning Performance. ETU-scores assigned for each interview step increased significantly from UT1 to UT2 for the introduction, eliciting, and planning phases (*U*-tests: all $z \geq 7.9$, all $p<.001$, see also Fig. 3, right hand side). For UT2, also scores for the interview skills empathy, communication, eliciting information, summarising, and transition were available for both scenarios. Averaged student scores (N=152) for the different skills ranged between 61.4 (SD=18.5) for transition and 89.7 (SD=15.5) for eliciting information with a mean of 77.0 (SD=11.4). Considering a possible score maximum of 100, students' objective performance was rather high, which is also reflected in the section scores and subjective feelings of success (NASA-TLX performance) reported above.

Motivation and Affect. Motivation, measured by four questions regarding the motivation to learn about clinical interviewing, to improve one's interview skills, and to apply them to real situation, remained at a high level (on 4-pt. rating scales all $Md \geq 3$). To account for the added affective scaffolding service in UT2 students have explicitly been asked about their *affective states*. The overall mean score derived from seven subscales of the TASS (M=64.4, SD=16.4) shows that students were in a positive affective state regarding different dimensions like mood, motivation or thinking activity after they had finished their training in the simulator.
The *Smiley Based Affect Indicator* (SBAI) was displayed 352 times to 158 unique users, i.e. between 1 and 8 times per user (M=2.23, SD=1.36, Md=2). Only one learner, however, actually made use of the affect report to indicate her current emotional state, while all other learners did not. It is not clear, though, what the reasons for not using the SBAI were – this will need further investigation by consulting students. As learners did not provide any Twitter IDs in the simulation, sentiment analysis of tweet streams to derive affect information was not possible.

Social Network Use, User Privacy and Trust. From 95 participants completing the cohort characterization survey, 81% use Facebook, 20% Twitter, and less than 5% LinkedIn, Flickr or MySpace. Asked for their social network IDs, only 13 students (13.7%) provided their usernames on different social media (mainly Facebook). On 4pt.-rating scales (from totally fine to will not provide) students also indicated that they are rather 'nervous' about providing their IDs or that they would not do it (Md=4, M=3.12, SD=1.12). Open responses from 60 students on reasons why they would/would not provide their IDs concern mainly privacy issues, such as use for friends and family only (25 entries), followed by another group of comments (13 entries) indicating insecurity and a lack of trust about what happens with their information. These results were underlined by participants' behaviour during their training with the simulator, where nobody provided a Twitter or Flickr ID, such that user model enrichment via the U-Sem service was not enabled and students could not experience additional adaptive coaching based on their enriched user model. Of course, with only 19 Twitter and one Flickr user (as per cohort survey), the number of users who could have potentially provided their IDs was very low in the first place. To investigate the correctness of user model augmentation, for the social IDs collected in the cohort survey a comparison of information derived via U-Sem services with explicitly queried or available learner data was attempted. The Twitter-based interest profile [15] could be derived for two users, confirming their medical background and studies: for one user a quite narrow set of interest topics could be identified, with one third clearly associated with health-related aspects; for the second user a slightly wider interest profile resulted, including but not limited to health and education related topics. A comparison of information from location detection [13] with cohort survey data was not possible due to the unavailability of Flickr user IDs.

4 Discussion and Conclusion

We presented an evaluation study of the improved ETU simulator with enriched user modelling and affective metacognitive scaffolding. The study was a follow up to a first user trial with metacognitive scaffolding only. This allowed a more specific evaluation of the services, because differences could be directly attributed to the two stages of development. Each type of advancement led to different research questions, which we want to take up again now.

Simulator Augmentation. The RolePlay Simulation Platform used for clinical interview training was enhanced with ImREAL services via U-Sem and AMSS. The extended simulator was evaluated with respect to service integration and perceived relevance of the learning experience. Relevance ratings did not increase significantly between the two trials, but users consider the training definitely as a relevant preparation for interviews with real patients. The simulation embeds an authentic scenario relevant to the real-world application context, as called for by adult learning theory [1],[2]. Technical flow ratings increased from the last user trial, which indicates a good integration of the services in the simulator and a smooth interplay of software components, thus allowing a fluent interaction with the simulation.

Affective Metacognitive Scaffolding Services. The already existing MSS was enhanced by adding an affective element to prompts. Also, the selection of prompts was modified by giving higher priorities to instructional designer weightings and taking into account possible affective impacts of prompts. That is, each of the scaffolding prompts was examined and rated as to whether it would have a neutral, negative or positive impact. These ratings were added to the rulebase of the scaffolding intervention selector. The simulation designer now also creates the mapping and alignment of metacognitive prompts to simulation steps. Furthermore, the SBAI was integrated into the feedback interface of the simulation. Research questions concerned the aspects SRL, learning experience, learning performance, motivation and affect. Effort and performance scores as well as students' reflection notes show that students are supported in their SRL activities. Nevertheless, students still find it difficult to estimate how well they will perform, whereby under- as well as over-estimations occurred.

With the improved service, the notes students take are increasingly of reflective nature, which points to an additional benefit that users get out of the enhanced simulation. However, the number of unique users who actually worked with the thinking prompts is rather low: 30 students left entries via AMSS and/or the ETU notepad. One important reason is that the ImREAL services are only available in practice mode, but almost 60% of the students did the training only in the assess mode (which is prerequisite for the practice mode). Thus, for future applications, it is necessary to find ways to attract more students into the practice mode and to actually use the prompts. Similar to this, the reluctance to use the SBAI needs to be further investigated. Potential reasons might be unawareness of this self-report tool, but also that students did not see any added value of using it. In the second case, more detailed information about the benefit for their learning experience (e.g. selection of prompts adapted to a user's current affective state) could be a possible solution.

With regard to learning experience and performance, results indicate that the enhancement of the simulator led to an improved learning experience for the students. In UT2 prompts were generally perceived as helpful. Furthermore, students from UT2 rated the overall feeling of flow higher than in UT1. Thus, the improved simulator was able to convey a stronger feeling of task involvement and fluency. With regard to workload, frustration went down and the feeling to be successful at a given task (performance) increased from UT1 to UT2. High performance is also confirmed by the objective scores assigned by the ETU simulator, where students obtained an average of 77 (out of 100). Their performance had increased from UT1 in the introduction, eliciting, and planning phase of the interviews. Motivation and affect measures showed that students had a high motivation regarding their learning task and that they were in positive affective states after the training.

Enriching User Models based on Digital Traces. The main idea behind U-Sem is to extract and structure relevant information from the social Web services users are active in and to use this information for enriching user models from client applications. These applications can consequently update their user profiles according to the newly received information and provide adaptive services without having the users to take initial assessments or fill out long surveys on their preferences. The Web services utilised by U-Sem for mining users' digital traces are Twitter and Flickr. Natural requirements for this procedure are users who actively use social platforms and the willingness of users to provide their social network IDs. There are two main points we can derive from this study. First, the use of social networks other than Facebook is not as widely spread as we initially assumed. Especially for Flickr, which is a main source for the digital traces used by U-Sem, only one participant indicated use of this platform. Second, the type of data used by U-Sem is publicly available on the Web and participants indicated that they believe that social networks are rather open (*Md* = 3.5 on a 4-pt. scale from 1-totally private to 4-totally open). Nevertheless, they were very reluctant to provide their social network IDs. Main reasons given were to maintain privacy and that the people behind the service are not known. Thus, for future research, these two aspects need to be taken into account by: (a) considering the type of sample (which might influence the type and extent of social networks used) and (b) by fostering trust in the service in order to prompt users to provide their user ID (e.g. by giving detailed explanation on the type of information that is used and on the way the information is retrieved, or by providing more information about the research group behind the service). By increasing users' willingness to share their user IDs, the correctness of user information derived from digital traces can be investigated in more detail in a next step.

Conclusion. This paper reports on further augmentations to the ETU RolePlay Simulation Platform with ImREAL services. Whilst further work needs to be performed to investigate outstanding issues of unwillingness to engage in affective reporting and supplying social IDs, we have shown that these augmentations further enhance the flow and fidelity of the simulation, leaving learners more motivated, engaged and competent. We finish with a quote from a user: 'I really like the simulation: it is very interesting and is directly associated to a real world problem.'

Acknowledgments. The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement no 257831 (ImREAL project) and could not be realised without the close collaboration between all ImREAL partners.

References

1. Knowles, M.: Andragogy in action. Jossey-Bass, San Francisco (1984)
2. Swartout, W.: Lessons learned from virtual humans. AI Magazine 31, 9–20 (2010)
3. Hetzner, S., Steiner, C., Dimitrova, V., Brna, P., Conlan, O.: Adult self-regulated learning through linking experience in simulated and real world: A holistic approach. In: Kloos, C.D., Gillet, D., Garcia, R.M.C., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 166–180. Springer, Heidelberg (2011)
4. Zimmerman, B.J.: A social cognitive view of self-regulated academic learning. Journal of Educational Psychology 81, 329–339 (1989)
5. Flavell, J.H.: Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry. American Psychologist 34, 906–911 (1979)
6. Azevedo, R., Hadwin, A.F.: Scaffolding self-regulated learning and metacognition – Implications for the design of computer-based scaffolds. Instructional Science 33, 367–379 (2005)
7. Moore, A., Hetzner, S., Pannese, L., Steiner, C., Brna, P., Conlan, O.: Affective metacognitive scaffolding for the enhancement of experiential training for adult learners. In: Workshop on Awareness and Reflection in Learning Networks (ARNets11), EC-TEL 2011, Palermo, Italy (2011)
8. Berthold, M., Moore, A., Steiner, C.M., Gaffney, C., Dagger, D., Albert, D., Kelly, F., Donohoe, G., Power, G., Conlan, O.: An initial evaluation of metacognitive scaffolding for experiential training simulators. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D.,

Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 23–36. Springer, Heidelberg (2012)
9. Macarthur, V., Moore, A., Mulwa, C., Conlan, O.: Towards a cognitive model to support self-reflection: Emulating traits and tasks in higher order schemata. In: Workshop on Augmenting the Learning Experience with Collaborative Reflection, EC-TEL 2011, Palermo, Italy (2011)
10. Schraw, G., Sperling Dennison, R.: Assessing metacognitive awareness. Contemporary Educational Psychology 19, 460–475 (1994)
11. Pekrun, R.: The impact of emotions on learning and achievement: towards a theory of cognitive/motivational mediators. Applied Psychology 41, 359–376 (1992)
12. Moore, A., Ascolese, A.: Explicit tools for affective indicators. In: Workshop on Augmenting the Learning Experience with Collaborative Reflection, EC-TEL 2011, Palermo, Italy (2011)
13. Hauff, C., Houben, G.-J.: Placing images on the world map: A microblog-based enrichment approach. In: SIGIR 2012, pp. 691–700 (2012)
14. Hauff, C., Houben, G.-J.: Deriving knowledge profiles from twitter. In: Kloos, C.D., Gillet, D., Garcia, R.M.C., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 139–152. Springer, Heidelberg (2011)
15. Tao, K., Abel, F., Gao, Q., Houben, G.-J.: TUMS: Twitter-based user modeling service. In: García-Castro, R., Fensel, D., Antoniou, G. (eds.) ESWC 2011. LNCS, vol. 7117, pp. 269–283. Springer, Heidelberg (2012)
16. Gao, Q., Abel, F., Houben, G.-J., Yu, Y.: A Comparative Study of Users' Microblogging Behavior on Sina Weibo and Twitter. In: Masthoff, J., Mobasher, B., Desmarais, M.C., Nkambou, R. (eds.) UMAP 2012. LNCS, vol. 7379, pp. 88–101. Springer, Heidelberg (2012)
17. Abel, A., Celik, I., Hauff, C., Hollink, L., Houben, G.-J.: U-Sem: Semantic enrichment, user modeling and mining of usage data on the social web. In: International Workshop on Usage Analysis and the Web of Data, USEWOD 2011 (2011)
18. Fill Giordano, R., Litzenberger, M., Berthold, M.: On the Assessment of strategies in self-regulated learning (SRL) – Differences in adolescents of different age group and school type. Poster Presented at the 9. Tagung der Österreichischen Gesellschaft für Psychologie, Salzburg, April 8-10 (2010)
19. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, N.M.P.A. (ed.) Human Mental Workload, pp. 239–250. North Holland Press, Amsterdam (1988)
20. Rheinberg, F., Vollmeyer, R., Engeser, S.: Die Erfassung des Flow-Erlebens [The assessment of flow]. In: Stiensmeier-Pelster, J., Rheinberg, F. (eds.) Diagnostik von Motivation und Selbstkonzept, pp. 261–279. Hogrefe, Göttingen (2003)
21. Nicolescu, A.B., Lulow, L.L., Ogden, C.A., Le-Niculescu, H., Salomon, D.R., Schork, N.J., Caligiuri, M.P., Lohr, J.B.: PhenoChipping of psychotic disorders: A novel approach for deconstructing and quantitating psychiatric phenotypes. American Journal of Medical Genetics Part B (Neuropsychiatric Genetics), 141B, 653–662 (2006)
22. Moore, A., Gaffney, C., Dagger, D., Kelly, F., Donohoe, G., Power, G., Conlan, O.: Preliminary content analysis of reflections from experiential training simulator support. In: Workshop on Self-Regulated Learning in Educational Technologies (SRL@ET), ITS 2012. Chania, Greece (2012)

Teaching Logic through Web-Based and Gamified Quizzing of Formal Arguments

Peter Øhrstrøm[1], Ulrik Sandborg-Petersen[1], Steinar Thorvaldsen[2], and Thomas Ploug[1]

[1] Department of Communication and Psychology, Aalborg University, 9000 Aalborg, Denmark
{poe,ulrikp,ploug}@hum.aau.dk

[2] Department of Education, University of Tromsø, 9037 Tromsø, Norway
steinar.thorvaldsen@uit.no

Abstract. In this paper the focus is on the construction, use, pedagogical potential, and long-term sustainability of certain web-based tools designed for teaching logic. A series of web-based tools have been implemented as a two-part system, and the tools have been tested and evaluated in a number of practical experiments. The first part of the system is a student-facing Java-Applet running in the student's browser, implemented using the Prolog programming language as embodied in a Java implementation called Prolog+CG. The second part is a teacher-oriented, server-based backend for logging the progress of students. In the paper, we provide a presentation of the pedagogical and technical ideas of construction that underpin the tools which have been made so far. It is explained how the tools can be developed as web applications using gamified quizzing. We then provide an evaluation of the potential of log data as learning analytics offered by these tools in the context of university courses introducing basic logic and formal aspects of argumentation. We also describe how we have used and evaluated the tools in a real learning context, using both quantitative, log-based data and qualitative interview. The evaluation also includes a discussion of the ethical aspects concerning the logging of student data. We also provide some insights on how the tools can and have been made sustainable.

1 Introduction

Since the rise of the medieval university in Europe the study of basic logic and argumentation has been seen as an essential component of basic academia. In the modern University, the various studies of computer science, information science and communication have made it even important to focus on logic and human reasoning. In consequence, it is commonplace at modern universities to offer basic courses in logic and argumentation. Such courses normally include the study of Aristotelian syllogistics and basic propositional arguments such as Modus Ponens and Modus Tollens (see [1] and [2]). On this background it seems

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 410–423, 2013.

obvious to look for tools which can support such courses and also tools which may support distance education in logic and argumentation.

A series of web-based tools for teaching logic have been implemented as two-part systems, and the tools have been tested and evaluated in a number of practical experiments. The first part of the system architecture which has been applied is a student-facing Java-Applet running in the student's browser, implemented using the Prolog programming language as embodied in a Java implementation called Prolog+CG (see [3] and [4]). The second part of the system architecture is a teacher-oriented, server-based backend for logging the progress of students and for locating the difficulties in the learning process. In [5] we have discussed these tools and their potential in logic teaching. This paper should be seen as a continuation of [5], and in the light of the long tradition within the area [9–13].

In section 2 of the present paper, a discussion can be found of the paedagogical and technical ideas of construction that underpin the tools which have been made so far, as well as a brief circumscription of the problem domain of teaching logic. In section 3 it is discussed how the tools designed for teaching logic can be developed as web applications using gamified quizzing. Section 4 contains a treatment of the teaching potential of these tools in the context of university courses introducing basic logic and formal aspects of argumentation, as well as a discussion of the ethical aspects concerning the logging of student data. This evaluation is based upon practical teaching experiences and various tests of the tools which have been carried out at Aalborg University, Denmark. Section 5 first investigates how the development of these kinds of tools for teaching logic can be carried out in a sustainable way. Section 6 offers some concluding remarks.

2 The Construction of Prolog+CG Tools for Logic Teaching

From a modern point of view classical syllogistics may be seen as a fragment of first order predicate calculus. A classical syllogism corresponds to an implication of the following kind:

$$(p \wedge q) \supset r$$

where each of the propositions p, q, and r matches one of the following four forms:

a(X, Y)	(read: "All X are Y")
e(X, Y)	(read: "No X are Y")
i(X, Y)	(read: "Some X are Y")
o(X, Y)	(read: "Some X are not Y")

These four functors were suggested by the medieval logicians referring to the vowels in the words "affirmo" (latin for "I confirm") and "nego" (latin for "I deny"), respectively. – The classical syllogisms occur in four different figures:

$$(x(M,P) \wedge y(S,M)) \supset z(S,P) \qquad (1^{\text{st}} \text{ figure})$$
$$(x(P,M) \wedge y(S,M)) \supset z(S,P) \qquad (2^{\text{nd}} \text{ figure})$$
$$(x(M,P) \wedge y(M,S)) \supset z(S,P) \qquad (3^{\text{rd}} \text{ figure})$$
$$(x(P,M) \wedge y(M,S)) \supset z(S,P) \qquad (4^{\text{th}} \text{ figure})$$

where x, y, $z \in \{$a, e, i, o$\}$ and where M, S, P are variables corresponding to "the middle term", "the subject" and "the predicate" (of the conclusion). In this way, 256 different syllogisms can be constructed. According to classical (Aristotelian) syllogistics, however, only 24 of them are valid. The medieval logicians named the valid syllogisms according to the vowels, {a, e, i, o}, involved. In this way the following artificial names were constructed (see [1]):

1$^{\text{st}}$ **figure:** barbara, celarent, darii, ferio, barbarix, feraxo
2$^{\text{nd}}$ **figure:** cesare, camestres, festino, baroco, camestrop, cesarox
3$^{\text{rd}}$ **figure:** darapti, disamis, datisi, felapton, bocardo, ferison
4$^{\text{th}}$ **figure:** bramantip, camenes, dimaris, fesapo, fresison, camenop

In these names the consonants signify the logical relations between the valid syllogisms, and they also indicate which rules of inference should be used in order to obtain the syllogism in question from syllogisms which were considered to be fundamental: barbara, celarent, darii, ferio. – In fact, the system of syllogisms may in this way be seen as the first axiomatic system ever (see [1] and [2]).

The Prolog+CG system dealing with syllogisms picks some terms from a given selection, an arbitrary figure number (1–4), and an arbitrary syllogism number (1–12). The syllogism numbers 1–6 point to the valid syllogisms listed above, whereas the numbers 7–12 stand for invalid syllogisms assumed to be somewhat "tempting":

1$^{\text{st}}$ **figure:** aia, oae, iai, ieo, iii, oao
2$^{\text{nd}}$ **figure:** oae, aoe, oio, ioo, ieo, oao
3$^{\text{rd}}$ **figure:** aaa, iaa, aia, eae, oae, eie
4$^{\text{th}}$ **figure:** aaa, aie, iaa, eae, eie, aee

In this way the system operates with a database of 48 syllogistic schemes among which 24 are invalid and 24 are valid according to Aristotelian syllogistics.

The Prolog+CG tools designed for teaching syllogistics make use of this database of syllogistic schemes. The steps in the program are the following:

1. A syllogistic scheme is picked at random.
2. The argument is presented to the user in terms of natural language with clear indication of the premises and the conclusion.
3. The user gives his evaluation of the validity of the argument.
4. The system states whether or not the answer given by the user is correct – if possible also providing a short explanation.

5. Unless certain criteria have been met, such as getting 5 or 10 correct answers in a row, the user is given the opportunity to go back to step 1 and thereby get another question to answer.

In addition to the tool dealing with syllogisms, we have implemented a tool dealing with basic propositional arguments. The valid arguments included in this tool can be proved using Modus Ponens, Modus Tollens, or usual disjunctive argumentation. This means that the following arguments are tested:

$$p \supset q, p \vdash q$$
$$p \supset q, \neg q \vdash \neg p$$
$$p \vee q, \neg p \vdash q$$
$$\neg(p \wedge q), p \vdash \neg q$$

Again the propositional tool is an implementation of the steps 1–5 mentioned above using variations of the propositional arguments mentioned. In this way a database of 36 schemes of propositional arguments has been constructed. 18 of these schemes are invalid, whereas the other 18 are valid.

The Prolog+CG tools concerning syllogisms and propositional arguments have been developed as web applications using gamified quizzing. The gamified versions of the tools count the number of correct answers in a row. When the user has got a certain fixed number of correct answers in a row, he or she has won the game. In the syllogistic tool and the propositional tool the winning criteria have been 5 and 10 correct answers in a row, respectively. All the student activity using the tools are logged in the teacher-oriented backend.

The reason these systems can be predicated with the term “gamified quizzing” is that they do not just present questions with concomitant required answers, as in a “bare” quiz. They also have a side-goal, namely getting 5 or 10 “points”, they use the concept of “streaks” (i.e., an unbroken sequence of correct answers), and they have a criterion for winning (and hence losing). Finally, external to the systems, the students were asked to attempt to “win” as fast as possible and let all the others know when they had finished. This last point gave a social dimension to the game which was added to the system by virtue of the co-location of the students in the same room, along with instructions from the teacher.

3 The Potential of the Prolog+CG Tools in a Teaching Context

As shown by [5] the teacher may obtain very useful information from the log-data of the student activities with the tools. In this way the teacher may also find out which types of arguments the students find easy and which they find difficult. This kind of log data seen as learning analytics can obviously be very useful for the teacher in his or her preparation and design of the lectures and

Table 1. Scores for syllogisms in figures 1, 2, and 3. In figure 1, no. 1-6 are valid, while no. 7-12 are invalid. In figure 2, no. 13-18 are valid, while no. 19-24 are invalid. In figure 3, no. 25-30 are valid, while no. 31-36 are invalid. In figure 4, no. 37-42 are valid, while no. 43-48 are invalid.

Syllogistic figure 1

Validity	No.	Score	No.	Score	No.	Score
Valid	1	0.82	2	0.64	3	0.92
Valid	4	0.77	5	0.49	6	0.51
Invalid	7	0.77	8	0.81	9	0.13
Invalid	10	0.39	11	0.14	12	0.19

Syllogistic figure 2

Validity	No.	Score	No.	Score	No.	Score
Valid	13	0.69	14	0.73	15	0.76
Valid	16	0.82	17	0.66	18	0.56
Invalid	19	0.61	20	0.80	21	0.28
Invalid	22	0.23	23	0.29	24	0.32

Syllogistic figure 3

Validity	No.	Score	No.	Score	No.	Score
Valid	25	0.67	26	0.85	27	0.88
Valid	28	0.61	29	0.82	30	0.81
Invalid	31	0.28	32	0.74	33	0.76
Invalid	34	0.35	35	0.78	36	0.66

Syllogistic figure 4

Validity	No.	Score	No.	Score	No.	Score
Valid	37	0.60	38	0.73	39	0.87
Valid	40	0.61	41	0.73	42	0.57
Invalid	43	0.28	44	0.87	45	0.62
Invalid	46	0.40	47	0.52	48	0.74

work in the classroom. Two interesting examples of results of this kind will be given here.

In continuation of the results obtained in [5] a study has been carried out with the syllogism-tool. During February 2012, an estimated total of 170 students at Aalborg University, Denmark, have been asked, alone or in small groups of up to 3 students in each, to work with a version of the syllogism tool which took them through all 48 schemes in a random order. They have been asked to do this within half an hour. A total of 103 groups participated, and this measurement was made before the students were introduced to syllogistics. The results found are depicted in Table 1.

Obviously, this measurement is very useful for teachers who want to run a course in syllogistic reasoning. Between sessions, the teacher can analyze the log data from the previous session, and then concentrate on the parts of the material which the students find particularly challenging. It is our experience that the system is very useful for this purpose.

The results listed above give important evidence for the difference between the syllogisms whose validity is drawn into question from a modern point of view (see Table 2(a)), and those which are not questioned from a modern perspective (see Table 2(b)). It is well known that the first group of 9 syllogisms are invalid in general if terms corresponding to empty sets are accepted, but valid if no empty-set terms are allowed in syllogisms.

On the basis of our investigation we may discuss how the students handle the 15 syllogisms that are not drawn into question within modern logic, versus the 9 Aristotelian syllogisms which have been questioned in modern logic. We may compare the group of these 15 versus the 9 questioned syllogisms and the results are shown in Table 3. The table shows strong statistical evidence against the

Table 2. The 24 Aristotelian syllogisms, divided into two groups based on whether they are drawn into question in modern logic or not. The reason to draw the 9 on the left into question is that they are invalid if terms denoting the empty set are allowed, but valid if empty set denotations are not allowed.

(a) 9 questioned syllogisms

Figure	Syllogisms
1st	barbarix, feraxo
2nd	camestrop, cesarox
3rd	darapti, felapton
4th	bramantip, fesapo, camenop

(b) 15 unquestioned syllogisms

Figure	Syllogisms
1st	barbara, celarent, darii, ferio
2nd	cesare, camestres, festino, baroco
3rd	disamis, datisi, bocardo, ferison
4th	camenes, dimaris, fresison

Table 3. This table summarizes how often the 103 student groups evaluated syllogisms in the two subgroups as valid. The difference between the evaluation of the 15 unquestioned syllogisms and the 9 questioned is strongly significant by the two sided chi square test.

	Correct reply? (Pre-test N=103 groups)	
Syllogism	Yes	No
Valid in modern syllogistics (15)	**1023**	**270**
Questioned modern syllogistics (9)	**454**	**319**
p-value	< 0.0001	

presumption that students will handle the two subgroups of syllogism equally well, and thus support our previous findings.

This study shows that although the students accept the Aristotelian notion of validity, seeing all 24 syllogisms as valid, they are much more certain of the validity when they are dealing with the 15 valid syllogisms than when they are dealing with the 9 syllogisms which are drawn into question if we allow for the use of terms corresponding to the empty set in the syllogisms.

Another interesting result based on log-data has to do with the propositional tool. Although only a few students (N=8) participated in the test of this tool it does in fact give strong evidence in support of the expectation that students handle modus ponens significantly better than modus tollens. The results can be summarized as in Table 4.

4 The Use of Gamified Quizzing in Teaching Logic

As mentioned in Section 2, the gamified version of the syllogistic tool will count the number of correct answers in a row, giving one point for each correct answer, and resetting the count to 0 when the user provides a wrong answer (thus ending a "streak"). When the user has got 5 correct answers in a row, he or she has won the game.

Table 4. This table summarizes counts of how often students (N=8) replied correctly to questions regarding the validity of propositional arguments. The difference between modus ponens and modus tollens is strongly significant by Fisher's exact test. (The chi-square test may not be used because some of the cell counts are low.)

	Correct reply? (N=8 students)	
	Yes	No
Modus ponens	**63**	**0**
Modus tollens	**21**	**55**
p-value	< 0.0001	

This gamified system was tested on March 7, 2013, at Aalborg University before and after a lecture on Aristotelian syllogistics. The students were asked to run the gamified syllogistic tool in groups of up to 4 both before and after their logic lecture, with an estimated 90 students participating. The students were asked to attempt to "win" the game as fast as possible, and were asked to indicate to all present in the room when they had won. All these test results were logged by the system. The score was computed as:

$$\text{Score} = \frac{\text{number of correct answers in last streak}}{\text{total answer count}}$$

The statistical analyses of the scoring data were performed using standard methods from descriptive statistics and statistical testing. The two sample t-test has been applied to detect increased score, and to look for significant differences between results from the Pre-test and the Post-test. The Pearson correlation coefficients were computed to look for association between time used at each task and scoring outcome. The chi-square test has been applied to detect differences in how students handle syllogisms with true or false conclusions.

The immediate experience during the Pre-test was that the students enjoyed the game and that the use of the game served as motivation for working with the theoretical question of validity which was introduced during the following lecture. After the late-afternoon lecture, a break was scheduled before the non-obligatory Post-test. Several students chose to go home during the break. It was, however, clear that the students who stayed to work with the game after the lecture now felt better equipped to deal with the questions of validity, and that they now had much more to talk about when facing the problems in the game.

4.1 Results

The Pre-test was performed with 53 groups of students. But only 33 of the groups achieved a score > 0, and hence the 20 groups with score=0 were not included in the analysis. It may be noted that this is standard practice when doing statistical calculations. The Post-test was performed with 24 groups of students. This big

Table 5. The table below summarizes the Pre-test versus Post-test results of the students score. The first table line is for the Pre-test versus Post-test of the student groups with score > 0, while the second line shows the same data for the 24 best groups. The first line supports statistical evidence against the presumption that students will handle syllogism equally well before and after the lecture. The p-values are by the two sample t-test (one sided not assuming equal variance).

	Mean Score	**SD Score**	**p-value vs. Post**
Pre (N=33)	0.455	0.261	0.040
Pre (N=24)	0.566	0.213	0.351
Post (N=24)	0.596	0.316	-

Table 6. The table summarizes the Pre-test versus Post-test results of the time which students spent on each task of the exercise. The first table line is for the Pre-test versus Post-test of the student groups with score > 0, while the second line shows the same data for the 24 best groups. Both support strong statistical evidence against the presumption that student will handle syllogism equally fast before and after the lecture. The p-values are by the two sample t-test (two sided not assuming equal variance).

	Median time (s)	**SD time (s)**	**p-value vs. Post**
Pre (N=33)	20.0	16.2	0.002
Pre (N=24)	20.6	15.1	0.005
Post (N=24)	34.3	27.0	-

dropout creates a bias for our analyses, and we had to compare the Post-test result versus both the 24 best results from the Pre-test as well as the 33 Pre-test results with better score than 0. Both results are shown in Table 5.

We also compared how much time the student groups spent thinking on each task of the syllogism exercise. Here we used the median time for each group, since such student groups may spend time talking about other things while they work. The results are shown in Table 6, and shows that the students spent significantly more time to think about each task after the lecture has been given.

Finally we looked at correlations between scoring and time spent on each task of the exercise. Table 7 presents the results, but shows no significant correlations, although they are positive.

4.2 Discussion of the Results

Results from Tables 5–8 shows mixed results. By learning classical syllogisms the students seem to learn to think, but not necessarily to achieve logically better results. However, Table 8 shows that the best groups base their estimation on validity of the syllogisms independent of the conclusion of the syllogism itself, while the Pre-test group of N=33 are initially more influenced in their logic by the closing conclusion of the syllogistic argument. In the last case they mainly base their evaluation of the validity of a syllogism on the final statement of the

Table 7. The Pearson correlation coefficients between score and time spent on each task of the syllogism exercise

	Correlation coefficient	**p-value**
Pre (N=33)	0.26	0.14
Pre (N=24)	0.03	0.88
Post (N=24)	0.31	0.14

Table 8. The table below summarizes counts of how often students in their evaluation of syllogisms agreed with a False or True statement in the conclusion of the syllogism. The first 2x2 subtable is for the Pre-test of the student groups with scores > 0, while the second 2x2 subtable shows the same data for the 24 best groups, and the third 2x2 subtable shows the results from the Post-test. The p-values are by the two sided chi-square test.

	Pre-test (N= 33 groups)		Pre-test (N= 24 best groups)		Post-test (N= 24 groups)	
	False	True	False	True	False	True
Statement in conclusion:	242	342	99	131	94	138
Student evaluation:	226	358	62	168	73	159
p-value:	0.37		0.0004		0.057	

syllogism alone, and how it is related to reality as they know it. The premises of the argument are to a large extent ignored.

4.3 Post-test Interviews: Results and Discussion

Following the lectures and the test, two semi-structured focus group interviews with a total of 11 students were conducted. The students volunteered to participate based on a general invitation. The interviews were structured around the following questions:

1. To what extent – if any – can the relevant system be of assistance in the acquisition of competence in classical logic?
2. What could or should be changed in the system in order to further the acquisition of competence in classical logic?
3. To what extent – if any – does the gaming element influence one's use of the system, i.e. one's play?
4. To what extent – if any – does the logging of results influence one's use of the system, i.e. one's play?

The answers to these four questions may be summarized as follows.

The system is generally considered to be useful for the attempt to acquire competence in classical logic (cf. 1 above). Several of the students summarize their experience in statements like “One gets a very fundamental sense of what logical validity is”, “You get this ‘aha experience’ about what logical validity is that you do not get in the lectures”, “You develop a sense of what it means for a statement to follow from another”, and “You catch a whiff of some underlying patterns and of logical validity”. Interestingly, in these reports the students refer to the acquisition of skills that seem intuitive and difficult to articulate. They speak about developing “a sense” having “an experience”, and catching “a whiff”. Several students contrast this intuitive grasp of logical validity with the theoretical understanding acquired in the lectures: “The system would be a fine supplement to the lectures because the lectures were theoretical in the sense that they provided a number of concepts to be used for theoretical analysis while the system provides a sense of logical validity in practice.” Another student remarks “I needed my notes from the lectures to aid me in creating some sort of overview of the syllogisms. It is easy to identify the figure ‘barbara’; from there on it gets harder.”

These reports point towards an important distinction in the skills and competences that may be acquired, on the one hand through the lectures, and on the other hand through the use of the system. In lectures, the students are taught concepts and models for the systematic analysis of the logical validity of arguments. In apprehending these concepts, the students first and foremost acquire the ability to apply the concepts to the systematic analysis of the logical validity of an argument. They do not first and foremost develop their ability to determine the validity of arguments without explicit, systematic analysis. By contrast the system builds the ability to “see” or “sense” the validity/invalidity without explicit, conceptual analysis.

It seems, however, that the ability to “see” or “sense” the validity/invalidity of an argument is something one may have independently of one’s ability to provide an account of what logical validity is and how it may be systematically determined. A great many people will thus be able to determine that the argument “If it rains, the road is wet. The road is wet. Therefore it rains,” is invalid on the grounds that the road may be wet for other reasons that it raining, without being able to define logical validity. Thus one might ask whether the system is, in fact, building competence in the conceptual analysis of logical validity, or whether it only reinforces pre-theoretical ability to do so.

It is, however, clear from the reports of the students that they do in fact apply the theoretical concepts from the lectures. As mentioned above, one student stated: “I needed my notes from the lectures to aid me in creating some sort of overview of the syllogisms. It is easy to identify the figure ‘barbara’ …”. The student thus explicitly identified a type of argument by means of the theoretical vocabulary, and would have wanted to analyze the arguments by means of the theoretical concepts, had they been available during the test. Two other students back this by stating that they to some extent “took notes on a piece of paper in order to determine the logical structure of the arguments.” Thus, in the mind of the students, there seems to be no clear separation of the pre-theoretical and the theoretical ability to determine the validity of arguments. By exercising one, one is exercising the other. At any rate, the analytical skills seem to be sharpened by doing rather than merely listening to a lecture.

In response to the second question of the interviews (cf. 2 above) on possible improvements of the system, the students suggest that “One should receive a notification if the validity of arguments of the same form are evaluated wrongly on several occasions,” and that “The syllogisms should also be shown in formalised and/or graphic form.” Both suggestions underscore the points made above. It is clear that the students to some extent perceive of the test as a test of the skills and competences acquired in the theoretical lectures, but also that the testing of these skills and competences could be more explicit. In terms of the specific suggestions for improvement of the system, they would clearly bridge this gap.

The students are in general reacting positively towards the gaming element in the test (cf. 3 above). Not surprisingly, the gaming element is noted by several to be motivating. It is recognised, however, that in this case – where the gaming element consists in the goal of getting 5 or 10 consecutive “right answers” – the gaming element may be especially appealing for the “stubborn, persistent and very competitive” students. The important lesson to be learned is straightforward. Different gaming elements attract different people. Future revisions and testing of the system should address this issue by more systematically incorporating and testing different gaming elements and their appeal to different types of “players”. This would obviously be primarily with the end-goal of improving the learning effect of the system on the students, though other ends could also be identified.

In response to the effects of the gaming element of the system, several of the students have observed an interesting pattern in their behaviour during the test. They report that “The gaming element made me inclined to take a chance at the point of having 4 consecutive right answers. After all, it was a game ....” The introduction of gaming elements thus seems to play a rather significant role not only for the students’ perception of the system, but also for their behaviour in using the system. A behavioural pattern of this kind clearly runs counter to the purpose of the system to motivate students to form some kind of reasoned judgement (pre-theoretical or theoretical) on the logical validity of arguments. It thus becomes of vital importance for future developments of the system to determine the extent to which the gaming elements trigger such behaviour.

The students all recognise that the logging of their answers had an effect on their behaviour even though they were anonymous (cf. 4 above). It is worth noting, however, that no one found the logging to be disturbing or caused him or her to be worried or anxious. On the contrary, all report that the logging initially had a positive effect on their motivation. As one says “It does mean something that you know someone is registering your score. It is positively motivating.” Several of the students note, however, that the actual effect of the logging on their results is not straightforward. Thus, in the course of the test the logging also had adverse effects on their motivation to do their best. “At the outset one is focused on evaluating the first five arguments correctly .... After having realised that it will not look good, you become more careless and start guessing." The exact effect on the scores generated by the logging is consequently hard to estimate.

5 Sustainability of the Tools

Sustainability of systems and tools within education has become a major issue in recent years, as people and projects may need to transcend the organizational, personnel, and budgetary boundaries of any one institution. One primary example is the Moodle[1] Course Management System, which has grown to become a world-wide phenomenon within education, with many contributors from several institutions and organizations in disparate locations around the world. The success of Moodle may arguably be attributed in part to the decision of its orginator, Martin Dougiamas, to license the source code of the system as Open Source. The effects of Open Sourcing software has been explored in various contexts, including [6] and [7].

Licensing the source code as Open Source by itself does not guarantee sustainability. Making the code generally available is obviously also necesssary. Good documentation is essential to supporting usability and the possibility of user adoption. Finally, building a community around the systems and tools in order to maintain the software beyond the capability, interest, and budgetary limits of the originators is of paramount importance for both sustainability and user adoption.

We have taken the first three steps towards sustainability of our tools listed above, of licensing as Open Source, making the source code available, and documenting it. The tools can be downloaded[2] from SourceForge.Net.

Whether the fourth step, namely building a community around the tools, will become a reality, remains to be seen.

6 Conclusion

We have in this paper presented two systems for teaching logic, namely the "syllog" system for teaching Aristotelian syllogistics, and the "proplog" system for teaching aspects of propositional logic. The two systems build on a number of common ideas, which may be formulated as imperatives. Though these imperatives are by no means normative for the broader range of educational software, they may illustrate some important aspects of the integration of software systems into traditional University-level courses on logic.

Support learning by doing: As [8] points out, higher learning outcomes can often be achieved through the student's doing an activity related to the theoretical material provided in lectures, than through lecturing along as a teaching

[1] `http://www.moodle.org/`

[2] `http://syllog.sourceforge.net/`

activity. In Section 4.3, we have seen how the students reported obtaining a deeper understanding of the theoretical material through using the system than they had acquired through the lectures. Thus supporting “learning by doing” is an essential educational feature of our tools.

Support learning of analytical inquiry skills: The process of analytically finding an answer to a question of validity in the field of logic may itself be founded upon the ability to ask relevant questions, especially since proofs may be seen as a sequence of propositions, the link between which consist of inference rules, validity arguments, application of axioms and theorems, etc. Thus the ability to think analytically about logical arguments involves the ability to inquire, or ask analytical questions of the problem at hand. The nature of the questions posed by our systems does not by itself ensure that the students will acquire skills in analytically oriented inquiry, but the repeated prompt to apply analytical skills, including methods of inquiry, may help instil in the student the basic potential for analytical inquiry. That is indeed a higher learning goal of many courses in logic, and our systems seem to support this goal, as indicated in Sections 3 and 4.3, as well as [5].

Provide learning incentives through gamification: Not all students will be equally inclined to take all learning activities seriously enough for their own good in meeting the learning outcome objectives of the courses which they attend. The element of gamification, as we have seen in 4.3, although eliciting mixed responses from the students, does seem to indicate that at least with the student focus group in question, the element of gamification had some positive effects on their motivation to learn. Gamification, of course, has many other ramifications and potentially positive outcomes than student motivation, including a strengthening of the element in the learning experience of the student of the application of the principle of repetition, which in turn may lead to higher learning outcomes. The proverbial abundant harvest founded upon diligent sowing applies equally well in farming as in many other kinds of work, including the work of learning. With our systems, the student is encouraged, through gamification, to sow more than they otherwise might have had the stamina or interest to do.

In future research, it would be interesting to:

- Expand the number of strategies for gamification. The current strategies could be enhanced by communicating streaks, percentages, and grade assessments at the end of each game.
- Expand the number of strategies for enhancing the learning experience of the student, for example by diagramming or otherwise showing the structure of the arguments.
- Elaborate on the kind of questions which the systems are able to ask of the students. In both of the current systems, the central question is the validity of the arguments. It might, however, be advantageous for the learning outcome of the student to be able to, say, pick a valid conclusion from a list of possible conclusions (valid or invalid), assess the truth-value of a given proposition within a given model, pick the correct figure and/or syllogism name for a given valid syllogism, and so on. Many different kinds of questions could be

devised, and it would be interesting to research the effect of the nature of the questions upon the learning outcomes and learning experiences of the students.

We have thus presented some systems for teaching and learning two kinds of logic, namely syllogistic logic and basic propositional logic. We have described their construction, and have evaluated their potential in teaching logic through both quantitative and qualitative methods. We have also provided some reflections on the sustainability of the tools, and have described how we have worked towards that goal by Open Sourcing the tools.

References

1. Parry, W.T., Hacker, E.A.: Aristotelian Logic. State University of New York Press (1991)
2. Aristotle: Prior Analytics. Jenkinson, A.J. (trans.) The Internet Classics Archive (1994-2000), `http://classics.mit.edu/Aristotle/prior.html`
3. Kabbaj, A., Janta-Polczynski, M.: From PROLOG++ to PROLOG+CG: A CG object-oriented logic programming language. In: Ganter, B., Mineau, G.W. (eds.) ICCS 2000. LNCS (LNAI), vol. 1867, pp. 540–554. Springer, Heidelberg (2000)
4. Kabbaj, A., Moulin, B., Gancet, J., Nadeau, D., Rouleau, O.: Uses, improvements, and extensions of Prolog+CG: Case studies. In: Delugach, H.S., Stumme, G. (eds.) ICCS 2001. LNCS (LNAI), vol. 2120, pp. 346–359. Springer, Heidelberg (2001)
5. Øhrstrøm, P., Sandborg-Petersen, U., Thorvaldsen, S., Ploug, T.: Classical Syllogisms in Logic Teaching. In: Pfeiffer, H.D., Ignatov, D.I., Poelmans, J., Gadiraju, N. (eds.) ICCS 2013. LNCS, vol. 7735, pp. 31–43. Springer, Heidelberg (2013)
6. Raymond, E.S.: The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. O'Reilly Media (2001)
7. Dougiamas, M., Taylor, P.C.: Moodle: Using Learning Communities to Create an Open Source Course Management System. In: Proceedings of EDMEDIA 2003 (2003)
8. Biggs, J.: What the Student Does: teaching for enhanced learning. Higher Education Research & Development 18(1) (1999)
9. Hoffman, B., McCrudden, M.T., Schraw, G., Hartley, K.: The Effects of Informational Complexity and Working Memory on Problem-Solving Efficiency. Asia Pacific Education Review 9(4), 464–474 (2008)
10. Bucciarelli, M., Johnson-Laird, P.N.: Strategies in syllogistic reasoning. Cognitive Science 23(3), 247–303 (1999)
11. Gilhooly, K.J., Logie, R.H., Wynn, V.: Syllogistic reasoning tasks, working memory, and skill. European Journal of Cognitive Psychology, 11(4), 473–498 (1999)
12. Johnson-Laird, P.N., Bara, B.G.: Syllogistic inference. Cognition 16, 1–61 (1984)
13. Schäfer, A., Holz, J., Leonhardt, T., Schroeder, U., Brauner, P., Ziefle, M.: From boring to scoring a collaborative serious game for learning and practicing mathematical logic for computer science education. Computer Science Education 23(2), 87–111 (2013)

The Student-Role in the One-to-One Computing Classroom: Tensions between Teacher-Centred Learning and Student-Centred Learning

Peter Bergström* and Stina Årebrand

Department of applied educational science, Umeå University, Umeå Sweden
peter.bergstrom@edusci.umu.se, stina.arebrand@hotmail.com

Abstract. One-to-one (1:1) computing has recently been scaled up and integrated into learning strategies, but there have been rather few studies about it so far. This explorative observation and interview study aims to gain increased understanding about the student role in the 1:1 computing classroom in upper secondary school. The results demonstrate a media-rich classroom based on four categories of affordances: students' note-taking; searching the Internet; social media; and laptops for duplication. The four categories of affordances delineate how teachers' behaviour is influencing students and their use of laptops in the designed learning activities. The designs of the 1:1 classrooms are based on technology-enhanced consumption of media as opposed to designs for technology-enhanced learning. It is concluded that the student role is diverse and stretched between principles of both teacher-centred learning and student-centred learning.

Keywords: one-to-one (1:1) computing, teachers behavior, constraints, affordances.

1 Introduction

The one-to-one (1:1) computing classroom is an environment based on one portable laptop per student, access to the Internet through wireless networks (WiFi), and a focus on using laptops in the school practice (e.g. for presentations and academic tasks) [14] (p. 331). Traditionally, schools have been organised based on principles of the analogue world, whereas 1:1 computing spans digital and analogue designs of the learning experience [7]. Goffman [9] used the concept of the façade to conceptualise both established roles and how settings construct roles based on the settings' design. For example, a young person, on walking into a school, takes on the formal role of student, and teachers and other students hold certain expectations related to that role. However, in the classroom, teachers' design of teaching and learning could either be teacher-centred or student-centred. Accordingly, the façade in settings of teacher-centred learning is based on hierarchical social relationships and knowledge that is transferred

* Corresponding author.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 424–429, 2013.

and remembered [1]. In contrast, the façade in settings of student-centred learning is characterised as holistic, active, and informal learning in diverse contexts [1] [13]. If the environment in which the role is played out changes, the role will probably be extended [11] to include additional façade [9]. Therefore, we need to know more about the student role based on how teachers' behaviour is influencing students and their use of laptops in the learning activities. To address this knowledge gap, we have asked the following two research questions: How can we understand the teachers' design of the 1:1 computing classroom? What expectations do the teachers hold about the student role in students' learning experience? To answer these questions we conducted an ethnographic study focusing on three 1:1 computing classrooms.

2 1:1 Computing in K-12 Education

Lei and Zhao [13] demonstrate four themes of how students use personal laptops: for specific learning tasks (e.g. note-taking, online searching), for communication (e.g. e-mail and instant messages), for expression (designing something), and for exploration (e.g. games and multimedia products). Dunleavy et al. [6] argue strongly that students' learning experience must be based on four design principles: learner-, assessment-, community-, and knowledge-centred (p. 451). An important aspect of the design is that a power relationship is created between, for example, Internet resources and textbooks, and one of them becomes dominant [3].

2.1 Theoretical Framwork

In this paper we argue that social actions in 1:1 computing classrooms affect the student role. Implicitly, the social actions are understood as being in relation to the design of the environment and the properties of the environment. Greeno [10] has developed Gibson's [8] theory of affordances from an ecological view towards a philosophical view by discussing affordances from Barwise's [2] situation theory. Greeno's [10] theory provides analytical tools for understanding how situations relate to other situations through the concept of conditional constraints (p. 339). Conditional constraints address regularities that provide a framework to outline situations from which affordances can be interpreted. Greeno's theoretical contribution highlights that if there are conditional constraints there are affordances that, by definition, can be both recognisable and perceived directly as properties of a design. For example, Chan et al. [5] state that 1:1 technology-enhanced learning (TEL) means that each student uses at least one computing device (p. 7), which brings out the following conditional constraint:

```
<<1:1 TEL>> --> <<students use at least one computing device>>
```

The above constraint highlights when the situation of 1:1 TEL means a relation to another situation where each student uses a personal computing device. Accordingly, the framework of conditional constraints and affordances will be applied here to students' activities involving 1:1 computing.

3 Method: Context, Data Collection, and Analysis

The study took place at an upper secondary school in the northern part of Sweden during the spring of 2012. The school has about 2,000 students and was chosen because of the municipality's 1:1 computing initiative. Each student and teacher was equipped with a laptop computer, and the classrooms were equipped with digital projectors, an interactive board (Smartboard), and a whiteboard. The whole school had WiFi access to the Internet and to a learning management system (LMS) called Fronter. Three social studies classes (A, B, and C) were chosen; there were 20 to 27 students in each class. We did not study students using laptops in the cafeteria or in their spare time.

We did 17 observations during three weeks of study. The observer sat at the back of the classroom and the observations were combined with informal conversations and interviews. Field notes were developed, as close as possible to the observations, into thick descriptions [12] (p. 153), which are characterised as narrative, describing, analysing, and interpreted.

Students were selected for focus group interviews based on the classroom observations where students were noted to be particularly verbal. The interviews followed a semi-structured interview guide covering two themes: how students study during lessons, and students' use of digital technologies. The interviews were digitally recorded and transcribed verbatim.

We analysed the empirical material using Boyatzis's [4] framework for thematic analysis. The process of thematic analysis is based on indicated essences, signs, episodes, and contrastive thinking that emerged. The themes were tested between the two authors. Here, the second author created the first set of themes, and then the first author assessed these themes in relation to the empirical material. This process was followed by discussions until the authors reached agreement about the themes.

4 Results

4.1 The Students' Activities

The students used technology for both schoolwork and non-school activities. For schoolwork, the students used the laptops most frequently for word processing (N=12) for taking notes and writing assignments. The students' used (N=7) a search engine for searching and accessing information. In a majority (N=13) of the learning activities, the students consumed media either from the teachers' digital presentations or from videos. Thereafter, the students' work was often done in pairs, as they answered questions about the subject of study from the teachers' digital presentation or video. During all the observations, students used their laptops and mobile phones to conduct non-school activities such as accessing social media applications, reading blogs, and playing online games. The students' activities and use of digital technologies indicate when one situation is related to another situation. Based on the students' activities and digital technology use, the following conditional constraint [10] for the 1:1 computing classroom is constructed:

```
 <<the use of digital technologies>> -->
<<refers to the influence on students' learning activities>>
```

In the next step, the above conditional constraint will be constituted through a set of affordances raised by the related situations. Informants' names have been replaced with pseudonyms, and quotations are included.

4.2 Four Properties of 1:1 Computing

The Affordances of Note-Taking. Note-taking is a key activity for students while teachers are presenting. The affordances are based on situations when students use the word processing application for note-taking in relation to situations of students' activities. The situations that influence students' learning range from their use of notes as they prepare for written exams to their development of note-taking strategies. However, a specific note-taking strategy was characterised as using the exact words and expressions (duplication) as in the teachers' digital presentations.

The Affordances of Search Engines. The affordances of search engines demonstrate two episodes where the situation of using search engines influences students' learning. In the first episode, in class C, the students' task was to complete a short test about the central bank. This activity prompted the students to use Google. Expected or not by their teacher, Maria, the students used phrases from the test and found the test on the Internet. Two students, Eva and Maya, said: " if you used the exact phrases from the questions when using Google [Eva] you got hits and it was rather simple to find the right answer [Maya]. In another episode, the students used tools to collect and read a variety of text resources on the Internet, such as Google, Google Scholar, and an essay database (http://www.uppsatser.se). In their research the students used sources (e.g. research papers) that were at a more advanced text level than the information in their textbook.

The Affordances of Social Media Applications. The affordances of social media highlight students' use in both school and non-school activities. In the school activities, social media (Facebook) has the affordances of sharing information (N=1). This episode demonstrates how students use social media for sharing information that supports the learning process. It was observed how Sofia had difficulty finding the facts she needed for an assignment and how Elin, sitting behind her, gave her some support. Sofia said: Could you send me the link about the nurses? Elin answered: I'll send it on Facebook. As for non-school activities, in all classes, during the teachers' digital presentations or during other activities such as group work, students were observed to be checking social media applications, playing games, and reading blogs. However, the students' lack of focus was not due only to the Internet and laptops; it was also observed when students were using mobile phones or chatting with each other.

The Affordances of Laptops for Duplication. In one episode, Maria taught by using the interactive board and a pencil to illustrate, step by step, how money is transferred in the economic system. Maria suggested that the students draw a picture similar to hers on their laptops. In parallel with Maria's teaching, the students started personal and collaborative sub-activities with a view to duplicating Marias drawing on their laptops. The sub-activities reflected students' diverse use of technology in one situation that related to supporting the students' learning when they were drawing the picture. Three affordances are indicated from the sub-activities. One affordance was highlighted when the students' Googled the subject of the drawing and found the picture. Another affordance of 1:1 computing was indicated when students suggested using the word processing application (Word) to write what Maria was explaining. In a third affordance, a group of students imitated Maria by trying to draw her picture using the word processing application.

5 Implications

The three 1:1 computing classrooms demonstrated greater emphasis on façade of teacher-centred learning based on technology-enhanced consumption (TEC) of media, and minor implications of façade of student-centred TEL. Thus, the teachers' behaviour and TEL designs are based on situations that refer to situations where students duplicate the teacher's digital presentation. The results indicate that the affordances of laptops support duplication by using text. Tablets would probably give richer affordances of taking notes through the affordances of the camera or of digital pens used on the tablets to make simple drawings. However, the implications of student-centred TEL were demonstrated when students shared links through social media applications, or when they used search engines. Teachers' TEL designs would benefit by integrating the social media applications that students use in non-school activities. To change practices towards student-centred TEL, conditional constraints [10] of the 1:1 computing classroom should be based on situations for using social media applications that refer to situations that strongly integrate students' preferences into the learning activities. Such designs could manifest collaborative note-taking activities and group reflections for deeper learning where all students have the possibility to contribute to a shared understanding.

Conclusion. This small ethnographic study confirms what other studies [3, 13] have found: that 1:1 computing classrooms are based on TEC designs for teacher-centred learning. From the teachers' perspective, expectations about the student role are based on traditional assumptions of TEC. Students' non-school activities make visible the tension between teacher-centred learning and student-centred learning, or in other words, the tension between façade of TEC and TEL designs.

References

1. Barr, R.B., Tagg, J.: From teaching to learning - A new paradigm for undergraduate education. Change (November/December 1995); Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
2. Barwise, J.: Constraints, channels and the flow of information. In: Aczel, P., Israel, D., Katagiri, Y., Peter, S. (eds.) Situation Theory and its Applications, vol. 3, pp. 3–27. Centre for the Study of Language Information, Stanford (1993)
3. Blikstad-Balas, M.: Digital literacy in upper secondary school: What do students use their laptops for during teacher instruction? Nordic Journal of Digital Literacy 7(2), 81–96 (2012)
4. Boyatzis, R.: Thematic analysis and code development: Transforming qualitative information. Sage Publication, London and New Delhi (1998)
5. Chan, T.W., Roschelle, J., Hsi, S., Kinshuk, Sharples, M., Brown, T., Paton, C., Cherniavsky, J., Pea, R., Norris, C., Soloway, E., Balacheff, N., Scardamalia, M., Dillenbourg, P., Looi, C.-K., Milrad, M.: One-to-one technology-enhanced learning: An opportunity for global research collaboration. Research and Practice in Technology Enhanced Learning 1(1), 3–29 (2006)
6. Dunleavy, M., Dextert, S., Heinecket, W.F.: What added value does a 1:1 student laptop ratio bring to technology-supported teaching and learning? Journal of Computer Assisted Learning 23, 440–452 (2007)
7. Fleischer, H.: Towards a phenomenological understanding of Web 2.0 and knowledge formation. Education Inquiry 2(3), 537–539 (2011)
8. Gibson, J.J.: The ecological approach to visual perception. Houghton Mifflin, Boston (1979)
9. Goffman, E.: The presentation of self in everyday life (In Swedish: Jaget och maskerna: en studie av vardagslivets dramatik) Bergstr/om, S.(trans.). Nordstedts Akademiska frlag, Stockholm (2009)
10. Greeno, J.G.: Gibson's affordances. Psychological Review 101(2), 336–342 (1994)
11. Jahnke, I.: Dynamics of social roles in a knowledge management community. Computers in Human Behavior 26, 533–546 (2010)
12. Kullberg. B.: Ethnography in the classroom (In Swedish: Etnografi i klassrummet) (2 ed.). Studentlitteratur, Lund (2004)
13. Lei, J., Zhao, Y.: One-to-one computing: What does it bring to schools? Journal of Educational Computing Research 39(2), 97–122 (2007)
14. Penuel, W.R.: Implementation and effects of one-to-one computing initiatives: A research synthesis. Journal of Research on Technology in Education 38(3), 329–349 (2006)

A Mobile Learning Community of Practice: Facilitating Conceptual Shifts in Pedagogy

Thomas Cochrane[1] and Laurent Antonczak[2]

[1] Centre for Learning And Teaching, AUT University, New Zealand
[2] School of Art & Design, Graphic Design, AUT University, New Zealand
{thomas.cochrane,laurent.antonczak}@aut.ac.nz

Abstract. University lecturers are usually experts within a specific context domain of knowledge, however they are seldom expert teachers or educational technology wizards. Facilitating professional development for lecturers to engage with innovation in teaching and learning utilizing new technologies is no simple process. However, we have found that structuring lecturer professional development around a sustained community of practice can result in a journey of conceptual and pedagogical transformation [1]. The increasing ubiquity of mobile devices and social media provides a platform for enabling pedagogical change. Thus we established a MOBile Community Of Practice (MOBCOP) of lecturers interested in researching the potential of mobile social media in education interested in exploring the concept of the pedagogy-andragogy-heutagogy continuum [2], and how mobile learning can be utilized as a catalyst [3] to move towards student-directed heutagogy. We argue that the MOPCOP framework is potentially applicable to a variety of higher education contexts.

Keywords: Communities of practice, mlearning, pedagogy, mobile, social media, heutagogy.

1 Introduction

Design education is often centred around students acquiring the skills and knowledge required to become active members of the design profession [4]. This can take the form of learning a range of core skills, such as the ability to make appropriate use of various types of media and specialised desktop editing software such as Photoshop, Final Cut Pro (FCP), and Dreamweaver. However, in a world increasingly dominated by mobile devices [5], the design parameters associated with small screen devices and their unique connectivity and collaboration affordances need to be part of the graphics design process. The context of the MOPCOP was a Bachelor of Graphics Design degree, with six lecturers participating in a community of practice investigating the potential impact of mobile social media within their courses. The impact of mobile devices and social media upon graphics design professionals' skills can be seen in the depreciation of Flash for html5, and the increasing first point of contact for web-based media being through users smartphones. Thus Graphics Design education needs to

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 430–435, 2013.

prepare students for this emerging environment. The MOPCOP experience traces participants' journeys from mobile social media technophobia to mobile social media evangelists. The activity of the MOPCOP was reified in the development of a range of mobile social media projects for students, and new forms of classroom interaction using mobile social media such as live video streaming, mobile eportfolios (for example Behance), and collaborative mobile video production. While the graphic design lecturers were already adept at utilizing desktop and laptop systems and software (such as Photoshop, FCP, and In Design) they had little experience of the design issues specific to mobile and social media and the link with their current practice. The MOBCOP (MOBile Community Of Practice) within the Graphics Design department was established to explore these important emergent issues.

Communities of practice (COP) is a social learning theory developed by Lave and Wenger [6] that has been extended by Wenger [7] and applied across many learning contexts. As an analytical framework for describing the process of becoming an active participant of a professional, social or learning community, the concept of COP resonates with the ethos of social media (participatory, social and user-generated). Wenger, McDermott and Snyder [8] have emphasized the importance of nurturing COPs, and the significant supportive and empowering role of technology and social media in communities of practice [9]. We therefore used a framework [1] that scaffolds the introduction of new technologies and new pedagogies in Graphic Design education by reframing lecturer professional development as a COP of like-minded peers supporting one-another exploring the potential of integrating mobile social media into the Graphic Design curriculum.

The ubiquity of mobile devices (particularly cameraphones and smartphones), alongside their connectivity, user content-creation, and collaborative affordances, has the potential to empower learners and disrupt traditional teacher-learner power relationships [10]. We argue that the disruptive nature of mobile devices can be harnessed to facilitate positive pedagogical change. Coupled with the user-centric and participatory nature of social media, mobile social media presents a unique opportunity for higher education to reconceptualize teacher and learning around authentic situated learning experiences and learner-generated content [2-3].

We have found the concept of the Pedagogy-Andragogy-Heutagogy (PAH) continuum to be a useful theoretical framework for measuring pedagogical change, from a focus upon teacher-directed to student-directed discovery within authentic contexts. This pedagogical change requires conceptual shifts regarding the roles of the teacher and the learner. As Luckin et al., [2] argue, a focus upon learner-directed activity need not be relegated to the sole domain of past-graduate studies, but can be used at any educational level via the careful design of learning activities and experiences. A heutagogical approach focuses upon higher-level learning that can produce graduates with creative and self-directed learning attributes that are critical to become active participants within the professional design community.

2 Research Methodology

The formation of a graphic design lecturer COP focused upon mobile media, in particular mobile film making guided by the following research questions:

- How can mobile social media be used as a catalyst for reconceptualising pedagogy, from teacher-directed paradigms to student-directed heutagogy using student-owned devices within the context of authentic collaborative projects.
- What is the potential for collaborative student-generated mobile video to create digital stories in an international team-based context.
- What are the affordances of mobile social media technologies to enable authentic learning contexts for Graphics Design.

A lecturer COP was established to explore the potential of mobile social media in graphic design education based upon our previously established framework [1]. The mobile COP was titled MOBCOP12. Six lecturers were supplied with an iPhone 4S, and committed to meeting together weekly to investigate mobile social media. The COP established the use of mobile social media tools for collaboration, including: Twitter, Wordpress, and Google+ (G+) (http://bit.ly/mobcop12) creating a bridge between their mobile and desktop/laptop workflows. Those COP members with previous mobile social media experience (including the authors as technology stewards) created an initial programme of mobile social media tools to investigate. Data collection and analysis were achieved by utilizing RSS feeds from participants' social media portfolios that they agreed to establish as part of participation within the COP. The participants were asked to regularly reflect upon their MOBCOP journey using YouTube, Vimeo, Wordpress blog posts, and to engage with Twitter, G+. Examples of the COP activity are curated at http://mobcop.wordpress.com. Emergent themes were identified and discussed during face-to-face workshop sessions with all the participants, who were tasked with creating projects for their students utilizing mobile social media for integration in the curriculum during the second semester of 2012.

3 Results

This section provides brief examples of the development of the MOBCOP through to the resultant implementation of mobile social media projects within the Graphics Design curriculum.

3.1 Establishing the Mobile Community of Practice (MOBCOP)

The goal of this project was to build capacity for innovation in teaching and learning within the department through the establishment and nurturing of a community of practice of lecturers (COP) in partnership with the universities Centre for Learning And Teaching (CfLAT) and also nationally / internationally. This COP investigated innovation in teaching and learning using social and mobile technologies within an action research framework. The project leveraged student-owned smartphones for student-generated content and student-generated contexts, and explored a variety of mobile media production and sharing tools.

Key to pedagogical innovation enabled by technology was developing lecturer competency with the technology so the project provided iPhones for participant lecturers to be used to develop their own eportfolios and research, so that they would become confident in integrating the use of mobile social media tools with their students from Semester 2 2012.

3.2 MOBCOP Outcomes

The MOBCOP was a journey for most participants from initial peripheral participation and limited engagement with mobile social media to active participation in an authentic community of practice. This was approached in a structured way, through a series of workshops on mobile social media topics held over a period of 7 weeks. These workshops, which were streamed using Bambuser (http://bambuser.com/channel/mobcop), exposed staff to new graphics design media in ways that enabled them to integrate it into their own teaching. Participants also began to use their mobile devices in research and development of a professional eportfolio (two members of the team). The activity of the MOBCOP was reified in the publication and sharing of peer reviewed research outputs based upon their experience and the impact on their students. This led to the development of new assessment and learning activities enabling student-generated content and contexts via student-owned mobile and social media tools, for example G+. Participants wrote an initial mobile social media proposal for implementation with their students in semester two 2012 at the beginning of the COP. This was then rewritten at the end of the first semester following their MOBCOP12 experience. These project outlines were uploaded and shared as Google Docs and provided a reified reflection activity evidencing significant development of the lecturers' understandings of the affordances of mobile social media for education and pedagogical change.

One MOBCOP participant was initially quite reticent to engage with mobile social media, and struggled to see the relevance for his teaching and his students' learning. However, after experimenting with the affordances of the iPhone and mobile social media throughout semester one, this lecturer implemented several mlearning projects within the courses that he taught, and by the end of the year had become a 'digital ninja'.

The benefits of this project were to understand the big picture of the device, the apps that can work from it, and how these can cause beneficial change in multiple areas of our program including different kinds of digital publishing over the next few years. (Lecturer blog post, 2012)

This lecturer created an engaging video summary report of the impact of the MOBCOP on their teaching practice, which can be seen at: https://vimeo.com/53726227.

Another lecturer explicitly implemented the PAH continuum within studio practice with a class group of eight Graphic Design students, level 7 (final year of their Bachelor Degree). Students were expected to use a variety of mobile social media tools for their course including: Youtube, G+, Bambuser, and create an eportfolio using Behance – a social network specialising in portfolios for design professionals http://behance.net. All the students were asked to deliver at least 3 mobile videos as a reflective statement as part of their digital workbook, but almost 50% of them did not meet these requirements – this was quite a departure from their previous learning experience that was much more teacher-directed. As a result of student engagement with mobile social media student's reflective statements exhibited a higher level of self-critical skills than in previous assessment approaches. Using mobile technology combined with mobile social media, specifically G+, facilitated students' conceptual shifts in sharing their work in progress and learning from it. Previously students were often reluctant to show their early sketches and didn't like to publicly (even in the class) discuss their ideas. G+ has been a powerful tool to deconstruct this reluctance. For example, one student who, prior to the project, barely engaged in class and in critique, in 2012 became one of the earliest adopters and users of mobile social media. In one example a student's self-analysis of her video project generated a discussion from student peers commenting on her progress via G+. In this case G+ was used to create a temporary community of practice amongst the students themselves. Thus the students took on board the COP model from their lecturer, creating a significant shift in their conception of the role of a learner. Unfortunately, this student COP lasted only one semester, since the students left the University, this COP is no longer active.

4 Discussion

One of the major positive and indirect outcomes of MOBCOP is that it created a research focus on the scholarship of teaching and learning within the department. MOBCOP acted as a catalyst to facilitate learning as conversation between staff and students via technology [11], MOBCOP created a new bridge between generations: communicating with the students though their personal devices was challenging but rewarding for all the MOBCOP participants. MOBCOP acted as a great support for staff professional development providing motivation for staff to revisit, or to update, their own learning and teaching practice. This led to redefining research as 'applied research' around the scholarship of teaching and learning. Using m-learning and mobile social media also challenged students to: initially consider a range of research strategies (from all disciplines), and carry out their project from the informed perspective of the reflective practitioner, as active participants within an authentic community of practice. Thus we argue that the PAH continuum is well suited for creative people who are naturally blending different methodological approaches in mixed mobile and desktop environments. This specific methodological framework is particularly relevant for practice-led research projects. Lastly, another positive aspect of a refocus upon heutagogy associated with mobile social media is that assessment deadlines were negotiated between the lecturer and the students and amended via the class online schedule using Google Drive.

5 Conclusions

Our MOBCOP experience reflects the transformational potential of a sustained community of practice to facilitate significant pedagogical and conceptual shifts for educators. These conceptual shifts are evidenced in the reified activity of MOBCOP resulting in informing new pedagogical strategies within the Graphics Design curriculum. We argue that the MOBCOP mobile social media framework is potentially transferable to a range of educational contexts, and will test this in future iterations of the associated projects [1].

References

1. Cochrane, T., Bateman, R.: A mobile web 2.0 framework: Reconceptualizing teaching and learning. In: Repetto, M., Trentin, G. (eds.) Using Network and Mobile Technology to Bridge Formal and Informal Learning, pp. 57–92. Chandos Publishing, Oxford (2013)
2. Luckin, R., Clark, W., Garnett, F., Whitworth, A., Akass, J., Cook, J., Day, P., Ecclesfield, N., Hamilton, T., Robertson, J.: Learner-Generated Contexts: A Framework to Support the Effective Use of Technology for Learning. In: Lee, M., McLoughlin, C. (eds.) Web 2.0-Based E-Learning: Applying Social Informatics for Tertiary Teaching, pp. 70–84. IGI Global, Hershey (2010)
3. Kukulska-Hulme, A.: Mobile learning as a catalyst for change. Open Learning: The Journal of Open and Distance Learning 25, 181–185 (2010)
4. Wikipedia, http://en.wikipedia.org/w/index.php?title=Atelier_Method&oldid=243469066
5. ITU, http://www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom.html
6. Lave, J., Wenger, E.: Situated Learning: Legitimate peripheral participation. Cambridge University Press, Cambridge (1991)
7. Wenger, E.: Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press, Cambridge (1998)
8. Wenger, E., McDermott, R., Snyder, W.: Cultivating communities of practice: a guide to managing knowledge. Harvard Business School Press, Boston (2002)
9. Wenger, E., White, N., Smith, J.: Digital Habitats: stewarding technology for communities. CPsquare, Portland (2009)
10. Sharples, M., Milrad, M., Arnedillo-Sanchez, I., Vavoula, G.: Mobile learning: Small devices, big issues. In: Balacheff, N., Ludvigsen, S., de Jong, T., Lazonder, A., Barnes, S., Montandon, L. (eds.) Technology Enhanced Learning: Principles and Products, pp. 233–249. Springer, Berlin (2009)
11. Laurillard, D.: Pedagogcal forms of mobile learning: framing research questions. In: Pachler, N. (ed.) Mobile Learning: Towards a Research Agenda, pp. 33–54. WLE Centre, Institute of Education, London (2007)

Using Smartphones and QR Codes for Supporting Students in Exploring Tree Species

Johan Eliasson, Ola Knutsson, Robert Ramberg, and Teresa-Cerratto Pargman

Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden
{je,knutsson,robban,tessy}@dsv.su.se

Abstract. Smartphones are increasingly being used on field trips to support students in exploring the natural world. In this paper we present a design and analysis of an inquiry-based learning field trip for primary school students. One problem for design is how to make use of smartphones to support, rather than distract, students in interacting with the physical environment. We approach this problem by comparing two alternative designs, where students use smartphones for identifying tree species either by using an identification instrument or by reading a text description. The results show that students made use of the instrument for identification, QR codes, for identifying tree species and made use of the text descriptions for searching for tree species. In this sense, QR codes, connecting contextual information on smartphones to the physical environment, work as a learning tool that may be used for orienting students in their interaction with the physical environment.

Keywords: Human-Computer Interaction, Interaction Design, Mobile Learning, Inquiry-Based Learning.

1 Introduction

Mobile technologies can be used to enrich learning experiences that take place outside the classroom. For instance, smartphones present features that allow learners to easily access, manipulate and share information. Moreover, they facilitate learners the access to their own resources, apps, and the Internet. The use of smartphones also make it possible for students to take photos and video, manipulate and share these. Together with positioning technologies, the use of smartphones may also help students in identifying relevant targets for their own learning goals; learning about their surroundings, when the digital information is tied to physical location. Taken together, mobile technologies can be used to orient students' focus in outdoor lessons.

On a small scale, mobile technologies have been used on field trips to support students in exploring the natural world. A few examples from earlier research are support for bird and butterfly watching [1, 2], visiting a botanical garden [3] or learning about woodland ecology [4]. Students may also use smartphones as tools for surveying or probing the environment they are visiting [5]. However, research in this field has found that introducing smartphones in field trips may lead students to focus on the technology rather than on the physical environment [3, 6-8].

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 436–441, 2013.

The field trip reported on in this paper is part of a project called mVisible, aimed at making abstract relations in natural sciences visible, tangible and concrete by using mobile technology. The learning tasks targeted in mVisible are framed within a structured inquiry-based learning process, with certain learning goals. The specific task focused on in this paper is for students to identify tree species, by using smartphones. The particular design problem of this paper is how to make use of smartphones to support, rather than distract, students in interacting with the physical environment. We approach this problem by comparing student interaction with smartphones in one design with QR codes and one design without QR codes. The comparison does not focus on learning quality and performance, however the comparison of student interaction with smartphones may be relevant for designing learning tasks and activities to meet with learning goals. The research question that guides the study is: How can we use mobile technology to support students in interacting with the physical environment in outdoor lessons?

2 Related Work

In the research field of human-computer interaction, there have been a number of projects studying the problem of interacting with mobile technology while on the move. These studies commonly build on two assumptions; where the first assumption is that the problem lies in how to match cognitive capacities (like cognitive load and short term memory) to what capacities are required to interact with the mobile technology. The second assumption is that the solution is to be found in the design of the user interface of the smartphone, through 'minimal attention' interfaces or making use of other modalities and thus not just relying on vision [9-11].

Referring to a review of 102 research projects in mobile learning [12], Göth [13] argues that among the 38 projects where the learning activities are set in the physical context, technology is too dominating in 28 of them (70%), In these projects the smartphones require continuous attention and interaction from the students, leading them to focus more on the smartphones than intended. This review of related research work suggests that mobile technology as distraction rather than support is a problem that is also noted by other researchers. However, with one exception the solutions suggested are not elaborated. In the only exception, Göth, Frohberg, & Schwabe [7], the evaluation of the follow-up study [14] resulted in only marginal improvements.

3 Method

In a previous study presented by Eliasson, et al., [15], interaction with smartphones was evaluated. Results from the evaluation suggested a model of contextual human-technology interaction (see Fig. 1). Building on the model, we design one condition that uses QR codes and one condition that does not use QR codes. In line with Eliasson, et al., [15], our primary focus remains on interaction with mobile technology rather than on assessment of learning with mobile technology.

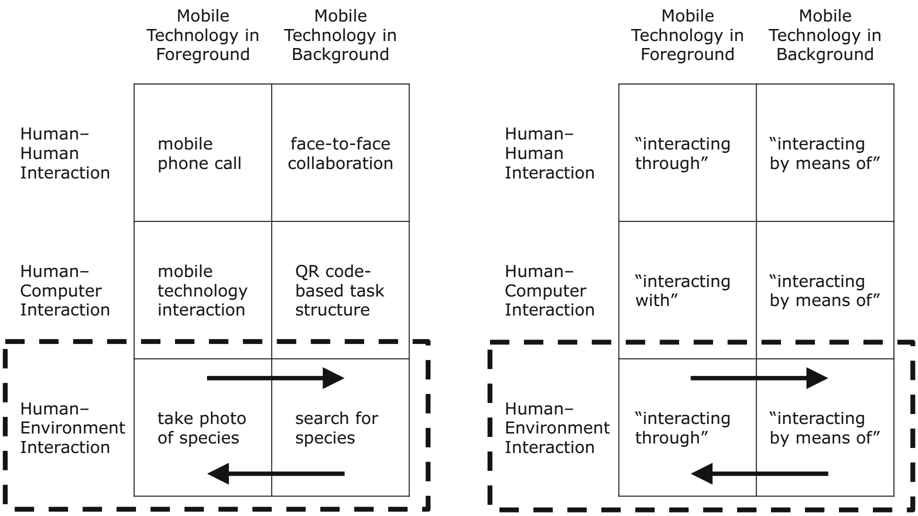


Fig. 1. Model of contextual human-technology interaction (left) and the model with modes of interaction applied (right). The dashed line marks the two squares focused on.

The problem focused on in this paper can be mapped to the bottom two squares of the model as can be seen in Fig. 1. Using the same argument as Buxton [16], the problem is how to design transitions where students can switch between interacting with mobile technology in the foreground and background of human-environment interaction. In our case this translates into how to design outdoor learning technology and lessons to orient students towards task-relevant features.

The model of contextual human-computer interaction must be operationalized to be meaningful in designing and analyzing specific learning designs. Arvola [17] express the design object of interaction design as interaction *with*, *through* and *by means of* the interactive system. We operationalize the bottom two squares of the model of contextual human-technology interaction as interacting *through* and interacting *by means of* mobile technology. One example of interacting through mobile technology is when students use smartphones for framing the physical environment when they take photos of characteristics of tree species (see Fig. 1). One example of interacting by means of mobile technology is when the task structure is given by smartphones and the students interact directly with the physical environment, however they do this in relation to mobile technology.

3.1 Study Setup

In mVisible, plants and trees are tagged with QR-codes, and when scanned with a smartphone the code gives additional information on the characteristics of each species. The inquiry-based learning activity starts in the field. Each student scans a QR code to initialize the smartphone to show a list of 5-7 available species in one of four different areas in the forest behind the school. Each group of three students then go to their "nature square", an area of 10x10 meters. They arrive at their nature square, which has been prepared with QR codes for each species and red and white tape markings in each corner. The task was designed to play out as follows: The students use their smartphones individually to scan QR codes attached to each species to identify them and take photos of what is characteristic of each species. Compared to the QR design, in the non-QR design the students identify the species by reading a description of the species on their smartphone and try to find a tree with the characteristics described in the text. Apart from the QR scanning functionality, the instructions given and the smartphone functionality were the same in the two designs.

3.2 Participants, Data Collection and Analysis

Seven groups of three students and two groups of two students, 25 in total, took part in the study. The students participated in the study as part of their biology curriculum. They were from two fourth grade classes at the same school. Apart from which students were known to be able to work together and which students were not, the teachers used high heterogeneity as the basis for group formations. Six out of nine groups had both female and male students. The teachers chose students for the groups to have similar differences in prior knowledge on the subject in all seven groups.

Data from six of seven groups of three students was used for analysis. The primary data used for analysis was video from handheld video cameras following each group in the field. In total 155 minutes of close-up video data was transcribed.

To compare how students interacted with the environment, we used two dimensions for measuring how the designs structure the learning activity in the QR design and the non-QR design: Number of photos and Time searching for the species.

Number of photos we use as an indication of to what extent students interact with the species and the characteristics of the species. From the data collected we use the number of photos each student took from the nature squares for this dimension. Time searching for the species we use as an indication of to what extent students interact with the physical environment where the species grow. From the video recordings we use the time the students need to go from one tree to the next for this dimension.

4 Discussion

The results are analyzed in relation to the model of contextual human-technology interaction presented in Fig. 1. The analysis shows that students searched for the right tree, corresponding to mobile technology in the background of human-environment interaction, more in the non-QR design than in the QR design (see Table 1). The analysis also shows that students took more photos, corresponding to mobile technology in the foreground of human-environment interaction, to a larger extent in the QR design than in the non-QR design.

In comparing the two designs we interpret that the students are mobile in and between contexts relevant for their learning goals to a larger extent in the QR design than they are in the non-QR design.

Table 1. Human-environment interaction: Number of photos and Time searching for the species

	Mobile Technology in Foreground	**Mobile Technology in Background**
Human-Environment Interaction	Number of photos: QR > Non-QR	Time searching for the species: Non-QR > QR

The results show that the students took more photos in the QR design than in the non-QR design. In taking photos the students interact *through* mobile technology. The results can then be interpreted as the students interact more *through* mobile technology in the QR design than in the non-QR design. In the model of contextual human-technology interaction, taking photos is in the bottom left square. This square corresponds to mobile technology in the foreground of human-environment interaction.

The results also show that the students searched more for the species in the non-QR design than in the QR design. In searching for the species the students interact *by means of* mobile technology. The results can then be interpreted as the students interact more *by means of* mobile technology in the non-QR design than in the QR design. In the model of contextual human-technology interaction, searching for the species is in the bottom right square. This square corresponds to mobile technology in the background of human-environment interaction.

Interacting with the environment through mobile technology means that students use smartphones as tools for interacting with the environment. Interacting with the environment through tools can be used in learning design for structuring the inquiry process. Finding the right tree is framed as scanning a QR code and identifying the characteristics of a tree is framed as capturing or representing aspects of the tree with the QR code in a few photos. Interacting with the environment by means of mobile technology means that students interact directly with the environment. The smartphones (together with rules, norms and specific instructions) give a structure by presenting tasks. Interacting directly with the physical environment can be used in learning design for allowing students to explore natural phenomena first-hand.

5 Conclusion

To recapitulate, the research question that guided the study was: How can we use mobile technology to support students in interacting with the physical environment in outdoor lessons? The findings indicate that; If one goal of the learning design is that students should use smartphones as a guide in exploring the natural world and to make abstract relations in natural science visible, then use QR codes, or similar technology for connecting digital and physical contexts. If one goal of the learning design is that students should explore the natural world more freely and establish relations between natural science and the physical environment themselves, then do not use QR codes, or similar technology for connecting digital and physical contexts.

References

1. Chen, Y.S., Kao, T.C., Sheu, J.P.: A mobile learning system for scaffolding bird watching learning. Journal of Computer Assisted Learning 19, 347–359 (2003)
2. Chen, Y.S., Kao, T.C., Yu, G.J., Sheu, J.P.: A mobile butterfly-watching learning system for supporting independent learning. In: WMTE 2004, pp. 11–18 (2004)
3. Naismith, L., Sharples, M., Ting, J.: Evaluation of CAERUS: A Context Aware Mobile Guide. In: 4th World Conference on mLearning: MLear 2005 (2005)
4. Vogel, B., Spikol, D., Kurti, A., Milrad, M.: Integrating mobile, web and sensory technologies to support inquiry-based science learning. In: 6th IEEE International Conference on Wireless, Mobile and Ubiquitous Technologies in Education, WMUTE 2010: Mobile Social Media for Learning and Education in Formal and Informal Settings, art. no. 5476539, pp. 65–72 (2010)
5. Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., Randell, C., Muller, H., O'Malley, C., Stanton, D., Thompson, M., Weal, M.: Ambient wood: Designing new forms of digital augmentation for learning outdoors. In: Proceedings of Interaction Design and Children, pp. 3–10 (2004)
6. Cole, H., Stanton, D.: Designing mobile technologies to support co-present collaboration. Personal and Ubiquitous Computing 7, 365–371 (2003)
7. Göth, C., Frohberg, D., Schwabe, G.: The focus problem in mobile learning. In: Fourth IEEE International Workshop on Wireless, Mobile and Ubiquitous Technology in Education, WMUTE 2006, pp. 153–160 (2006)
8. Cook, J.: Mobile Phones as Mediating Tools Within Augmented Contexts for Development. International Journal of Mobile and Blended Learning 2, 1–12 (2010)
9. Lumsden, J., Brewster, S.: A paradigm shift: Alternative interaction techniques for use with mobile & wearable devices. In: Proceedings of the 2003 Conference of the Centre for Advanced Studies on Collaborative Research, pp. 197–210 (2003)
10. Oulasvirta, A., Tamminen, S., Roto, V., Kuorelahti, J.: Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile HCI, pp. 919–928 (2005)
11. Pascoe, J., Ryan, N., Morse, D.: Using while moving: HCI issues in fieldwork environments. ACM Transactions on Computer-Human Interaction 7, 417–437 (2000)
12. Frohberg, D., Göth, C., Schwabe, G.: Mobile Learning projects - a critical analysis of the state of the art: Original article. Journal of Computer Assisted Learning 25, 307–331 (2009)
13. Göth, C.: Mobile Exploration: Lernen Im Physischen Kontext. PhD Thesis, Zürich University, Switzerland (2009)
14. Göth, C., Schwabe, G.: Navigation support for mobile learning. In: HICSS 2010 Proceedings of the 2010 43rd Hawaii International Conference on System Sciences (2010)
15. Eliasson, J., Knutsson, O., Nouri, J., Karlsson, O., Ramberg, R., Cerratto Pargman, T.: Evaluating Interaction with Mobile Devices on a Field Trip. In: Proceedings of the 2012 IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education (WMUTE 2012), pp. 92–96 (2012)
16. Buxton, B.: Integrating the periphery and context: A new model of telematics. In: Graphics Interface, pp. 239–246 (1995)
17. Arvola, M.: Shades of use: the dynamics of interaction design for sociable use. University: PhD Thesis, Linköping (2004)

Using Educational Domain Models for Automatic Item Generation Beyond Factual Knowledge Assessment

Muriel Foulonneau and Eric Ras

Public Research Centre Henri Tudor,
29, av. John F. Kennedy
L-1855 Luxembourg, Luxembourg
{muriel.foulonneau,eric.ras}@tudor.lu

Abstract. The Semantic Web offers many opportunities for reusing datasets and domain models in the field of education and assessment in particular. We have conducted research to generate test questions from Semantic resources. The reuse of semantic resources raises however challenges, since all data and models have not been conceived to be directly used for educational purposes. We have therefore analysed existing domain models created specifically for educational contexts to identify structures and relations of the model that can help deem the relevance of a particular domain model for automatic item generation. We present an initial set of conditions which can help identifying a relevant domain model to be used in an educational context. We also suggest a mechanism to relate them to levels of knowledge to be assessed in generated items.

Keywords: domain model, Semantic Web, assessment item, model pattern.

1 Introduction

More and more, domain models are created to represent phenomena, problems and issues on the Web. The Semantic Web [1] makes available to machines many pieces of knowledge which could be used as domain models (e.g., [2]). The educational sector uses domain models to generate learning and/or assessment resources (e.g., [2-4]). The ability to reuse semantic resources from the Web as domain models to support education is critical. Indeed, certain projects create domain models to support the generation of learning resources, such as APOSDLE or DYNALEARN. However, this requires gathering experts for a modelling exercise. In order to ensure the update of the model, the experts should even be gathered on a regular basis. This makes it very unlikely that the use of domain models can be generalized for educational domains in fast evolution (e.g., technologies), and specialized knowledge for professional environments where the number of learners is more limited and the curricula are not as formalized as for childrens' educational systems.

As a result, research has initiated the generation of assessment items (AIG) from the Semantic Web ([5-7]). By reusing existing domain models, the generation of assessment items can produce a wider variety of items. However, current approaches only address the lowest level of factual knowledge.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 442–447, 2013.

Our aim is to generate items from the Semantic Web to assess more dimensions of knowledge (e.g., conceptual knowledge). As a first step, we need to understand whether the Semantic Web includes relevant model patterns, which can be used for item generation. A long-term objective is to develop and evaluate a framework which helps relating model patterns to types of assessment items.

In this paper, we analyse the level of knowledge assessed by the items which are generated from the Semantic Web (Section 3), we analysed the model patterns presented in three projects (Section 4), and finally we draw perspectives on the ability to reuse models from the Semantic Web.

2 Existing Work on Models to Support Education and Assessment

Domain models have been created in education, in particular with ontologies. Most approaches involve text mining and/or domain and modelling experts to derive domain models [8]. Further, tools for the creation and validation of models by experts, such as Wikis [9] or question generation interfaces [10] are available. However, these works do not define the specificities of the domain models used for education.

Ying and Yang propose a mapping between modelling patterns and knowledge dimensions identified in the Bloom taxonomy [11]. They state that factual knowledge for instance is represented by holonymy and meronymy, conceptual knowledge by hypernymy, hyponymy, and cause/effect relations, finally procedural knowledge by formula, procedure, and condition relations. It defines the type of modelling patterns which would be suitable to represent the knowledge dimensions of the Bloom taxonomy. The way in which these relations are actually represented is not analysed (which properties are typically used) and it does not provide a link to assessment item types for instance.

In NLP approaches to the generation of questions from texts, Shah proposes a mapping between discourse connectives (e.g., *because*), relations (causal), and question types (why) [12]. This creates a framework to identify textual patterns and relate them to questions but not to knowledge levels.

The Bloom taxonomy is considered not to be useful for assessment because item and test performance are more related to knowledge and skills levels than to cognitive processes respectively cognitive demand [13]. Therefore, we decided to address only the knowledge dimension of the revised Bloom's taxonomy [14]. Anderson and Krathwohl prefer distinguishing knowledge about discrete concrete elements (i.e., between terms and facts) from larger, more organized aggregation of knowledge (i.e., general concepts, principles, models, or theories):

Factual knowledge refers to the basic elements that people use in communicating within their discipline and which can be isolated from their context. Anderson and Krathwohl refined factual knowledge into *knowledge of terminology*, which includes knowledge of verbal and nonverbal labels and symbols (e.g., words, numerals, pictures), and *knowledge of specific details and elements* (e.g., events, locations, people, dates, sources of information). *Conceptual knowledge* is more complex, organised knowledge. Several subtypes exist: *knowledge of classifications and* *categories* is more general and abstract than knowledge of terminology and specific facts. It connects specific elements by links. *Knowledge of principles and generalizations* is composed of classifications and categories and is used to study phenomena or solve problems in a discipline. Finally, *knowledge of theories, models, and structures,* which are the most abstract formulations of conceptual knowledge. The level of *procedural knowledge* and *metacognitive knowledge* are not addressed by the models analysed in this paper.

3 The Generation of Questions from the Semantic Web

We reuse the DBpedia semantic dataset to generate choice and match items in IMS-QTI format [2] and assign a knowledge level (Table 1). In Q1, we only request the property of an entity. In Q2, we use a relation between an entity (country) and a pictoral representation (flag). The selection of countries is however limited to instances of a particular class. Q3 use a property *successor* but with the definition of entities as instances of the *KingOfFrance* class. Q4 uses both a property of the country entity and a relation to an external Web resource to include a feedback. In Q5 we use a property *meshname* with a SKOS category as a range, i.e., a thesaurus.

Table 1. Knowledge level addressed by different questions

Q	Item stem	Ontology	Know. level
Q1	*What is the capital of [Azerbaijan]*	Infobox	F
Q2	*Which country is represented by this flag?*	Infobox, FOAF, YAGO	F
Q3	*Who succeeded to [Charles VII the Victorious] as ruler of France ?*	Infobox, DBpedia, ontology, YAGO	F
Q4	*What is the capital of [Argentina]?* With feedback	Infobox ,YAGO	F
Q5	*Which category does [Asthma] belong to?*	SKOS, Dublin Core, MESH, Infobox	C
Q6	*Which is the style of the following painting?*	Freebase	C

In Q6 we use a Freebase dataset to identify the artistic movement of a particular painting represented by its picture. The SPARQL query uses properties of an art work, although in DBpedia, the Guernica painting has a dcmi:subject property *http://dbpedia.org/page/Category:Cubist_paintings*. Its classification is therefore only represented as a property pointing to a taxonomy, instead of an instantiation or subclass relation.

In Q1 through Q4 we only created items to assess factual knowledge through properties. For Q5 and Q6, we assess conceptual knowledge, since we address a classification of knowledge through a taxonomy.

Álvaro Rey et al. have developed (multiple choice questions) using Linked Data (i.e., DBPedia) [7]. Using the RDF triple as input (example: “dbp:Napoleon dbpr:placeOfDeath dbp:Saint_Helena.“), the approach creates so-called fact statements (open question allowing natural language answers as response) and quizzes (multiple choice). The items address the level of factual knowledge only and the subtype of *knowledge of classification and categories* on the conceptual level.

Bratas et al. propose the Quiz Engine to create a complex game composed of multiple choice, matching and hangman item using facts generated from triples of DBpedia [6]. This approach tackles only the factual knowledge level.

4 Characteristics of Domain Models for Educational Purposes

We have analysed the characteristics of domain models in three cases:

Case 1: In the FP7 APOSDLE project, ontologies were created by domain experts with a wiki interface [8]. Kump describes domain models for learning with network models including task and skill related relations such as *prerequisite* and *issubtask-of* [15]. In Ghidini et al., a domain concept template proposes using a concept represented in a domain model with synonymy, meronymy, a description as well as a list of properties [8]. A *is_prerequisite_of* relation can be used to define an ordering between the concepts of a domain model in the learning path. The relation with education is provided by the creation of task models which relate to one or more concepts of the domain model.

Table 2. Modelling components identified in models of the DynaLearn project

	Set B [16]	Set A [17]	Set D [18]	Set C [19]
Number of topics	7	13	8	31
Quantity spaces	x	x	x	x
Rel. between quantity spaces	x	x		
Cause	x	x	x	x
Impact	x	x	x	x
Factors/influence	x	x	x	x
Events			x	x
Cond. Statements / knowledge		x	x	x
Conditional impact	x			
Feedback loop			x	x
Competing rates				x
Rate balance	x		x	
System patterns				x
Phenomenon			x	
Mechanism			x	
Combination of scales			x	
Process		x	x	
Taxonomies / entity hierarch.	x	x		
Cycle / cycle component		x		

Case 2: In the scope of the FP7 DYNALEARN project, six sets of models were created from which we analysed four. In Table 2, the sets are presented with the number of topics. Each topic can include multiple models. Borisova et al. distinguish between multiple types of models [16]. For each topic, they create a concept map which represents concepts and their relations, including influence, impact or meronymy for instance but no quantities or proportions, configurations relating the different entities, entity hierarchies, and a causal model. In addition, model fragments represent quantity spaces, relations between values of quantity spaces, impacts on

proportions or quantities, and influences. Noble also distinguishes between basic causal models, differential causal models where multiple factors can cause a phenomenon and causal state graph models [17]. DynaLearn sets of models are mainly dedicated to generate assessment resources on factual (*Give the amount of Value*, [4]), conceptual (*What is going to happen with Amount during simulation*?, [4]), and procedural knowledge (decision to remove macrophytes when the biomass reaches a certain level so as to avoid pollution in the phytoremediation model in [18]). The analysis of the model characteristics shows the importance of causal relations (subtype of conceptual knowledge in the knowledge taxonomy [11]), processes, conditions, and quantified impacts.

Case 3: In Djaghloul et al., we created an ontology from an interview with a dentistry professor on caries in order to generate assessment items [10]. Modelled by an ontology expert, it was then validated by the teacher to support the generation of assessment items. It includes instances, subclasses, meronymies, and different types of impact relations.

5 Conclusion: Towards Reusing Models from the Semantic Web

We propose using the Web as a source of knowledge to support education. In Foulonneau we have investigated the data quality issues which can prevent the creation of high quality assessment items [2]. In Foulonneau et al. we have proposed mechanisms to define the educational context in which sub models can be used to generate items [5]. In this paper, we aimed to investigate the limitations of the semantic resources published on the Web to be reused as educational resources.

The most commonly used vocabularies on the Linked Data Cloud [1] is Dublin Core. However, Dublin Core essentially aims to provide descriptive information about resources. A variety of relations can be used, such as *isPartOf*, *hasPart* (meronymy, factual knowledge) as well as the more general concept of *relation*, which can be used for comparisons for instance (conceptual knowledge). In addition, among the most common ontologies, a number of taxonomies and classifications are represented, including vu-wordnet, geospecies and uniprot.

The ability to generate assessment items on factual knowledge is a great opportunity. However, in most cases, AIG has been limited so far to the assessment of factual knowledge. This is due to some extent to the inherently descriptive nature of Semantic Web resources which describe concepts rather than processes, impacts, and factors. This makes it a relevant source of items assessing factual knowledge and to certain extent conceptual knowledge through Knowledge Organization Systems (e.g., SKOS). These however raise quality issues related to the reuse of taxonomies which have not been conceived for educational purposes (e.g., Obesity as a type of childhood illness in [2]). The development of vocabularies to represent events (e.g., the LODE ontology for Linking Open Descriptions of Events) or processes (e.g., the Open Provenance Model) can help enrich existing datasets published on the Web with relevant information and hence to support the generation of assessment items that go beyond the assessment of factual knowledge.

References

1. Bizer, C., Jentzsch, A., Cyganiak, R.: State of the LOD Cloud, Version 0.3, 09/19/2011 (2013), http://lod-cloud.net/state/ (retrieved April 8, 2013)
2. Foulonneau, M.: Generating Educational Assessment Items from Linked Open Data: the Case of DBpedia. In: García-Castro, R., Fensel, D., Antoniou, G. (eds.) ESWC 2011. LNCS, vol. 7117, pp. 16–27. Springer, Heidelberg (2012)
3. Papasalouros, A., Kotis, K., Kanaris, K.: Automatic generation of tests from domain and multimedia ontologies. Interactive Learning Environments 19, 5–23 (2011)
4. Linnebank, F., Liem, J., Bredeweg, B.: Question generation and answering. DynaLearn, Deliverable D3.3, EC FP7 STREP project 231526 (2010)
5. Foulonneau, M., Grouès, G.: Common vs. expert knowledge: Making the Semantic Web an educational model. In: Linked Learning 2012 at WWW Conference, Lyon, France (2012)
6. Bratsas, C., Chrysou, D.E., Eftychiadou, A., Kontokostas, D., Bamidis, P., Antoniou, I.: Semantic Web Game Based Learning: An I18n approach with Greek DBpedia. In: Linked Learning 2012 at WWW Conference, Lyon, France (2012)
7. Álvaro Rey, G., Celino, I., Damova, M., Damljanovic, D., Li, N., Alexopoulos, P., Devedzic, V.: Semi-automatic generation of quizzes and learning artifacts from Linked Data. In: Linked Learning 2012 at WWW Conference, Lyon, France (2012)
8. Ghidini, C., Rospocher, M., Kump, B., Pammer, V., Faatz, A., Zinnen, A.: Integrated Modelling Methodology - Version 2.0. - Deliverable D1.6. Aposdle, EC FP7 IP project 027023. Fondazione Bruno Kessler, Trento (2009)
9. Pammer, V.: Automatic Support for Ontology Evaluation - Review of Entailed Statements and Assertional Effects for OWL Ontologies. Faculty of Computer Science. PhD. Graz University of Technology, Graz (2010)
10. Djaghloul, Y., Foulonneau, M., Jadoul, R.: From informal knowledge to High-Quality e-Assessment process: The OAT Semantic Approach. In: 15th International Conference on Interactive Collaborative Learning and 41st International Conference on Engineering Pedagogy, Villach, Austria (2012)
11. Ying, M.-H., Yang, H.-L.: Computer-Aided Generation of Item Banks Based on Ontology and Bloom's Taxonomy. In: Li, F., Zhao, J., Shih, T.K., Lau, R., Li, Q., McLeod, D. (eds.) ICWL 2008. LNCS, vol. 5145, pp. 157–166. Springer, Heidelberg (2008)
12. Shah, R.: Automatic Question Generation using Discourse Cues and Distractor Selection for Cloze Questions. Language Technology and Research Center (LTRC), thesis. International Institute of Information Technology, Hyderabad (2012)
13. Haladyna, T.M.: Automatic Item Generation - A historical perspective. In: Gierl, M.J., Haladyna, T.M. (eds.) Automatic Item Generation. Routledge, New York (2013)
14. Anderson, L.W., Krathwohl, D.R.: A Taxonomy for Learning, Teaching, and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives. Longman, New York (2001)
15. Kump, B.: Evaluating the Domain Model of Adaptive Work-integrated Learning Systems. PhD dissertation. Karl-Franzens Universität Graz, Graz (2010)
16. Borisova, P., Nakova, E., Uzunov, Y.: Central Laboratory of General Ecology - Basic topics and models - Deliverable D6.2.3. DynaLearn, EC FP7 STREP project 231526. Central Laboratory of General Ecology, Sofia (2010)
17. Noble, R.: University of Hull – Basic Topics and Models - Deliverable D6.2.2. DynaLearn, EC FP7 STREP project 231526. University of Hull, Yorkshire (2010)
18. Salles, P., Souza, A., De Sá, I.G., Leite, G.M.F., Tunholi, V.: Costa e Silva, P.A.: FUB - Advanced topics and models - Deliverable D6.4.1. DynaLearn, EC FP7 STREP project 231526 (2011)
19. Zitek, A., Poppe, M., Stelzhammer, M., Jung, A., Zacharias, M., Muhar, S.: University of Natural Resources and Life Sciences, Vienna (BOKU) - Basic Topics and Models - Deliverable D6.2.5.: DynaLearn, EC FP7 STREP project 231526 (2010)

Towards an Integrated Learning Design Environment

Davinia Hernández-Leo[1], Jonathan Chacón[1],
Luis P. Prieto[2], Juan I. Asensio-Pérez[2], and Michael Derntl[3]

[1] ICT Department, Universitat Pompeu Fabra, Barcelona, Spain
{davinia.hernandez,jonathan.chacon}@upf.edu
[2] School of Telecommunication Engineering, University of Valladolid, Spain
lprisan@gsic.uva.es, juaase@tel.uva.es
[3] Advanced Community Information Systems, RWTH Aachen University, Germany
derntl@dbis.rwth-aachen.de

Abstract. Learning design research focuses on how educators can act as designers of technology-supported learning activities according to their specific educational needs and objectives. To foster and sustain the adoption of Learning design, the METIS project is working towards the implementation of an Integrated Learning Design Environment (ILDE). This paper presents the vision for the ILDE and how user requirements from three educational institutions in vocational training, higher and adult education justify the need for this vision. The paper discusses the data collected in questionnaires, on-line interviews and face-to-face group work with the end-users, as a first phase in the design-based research methodology applied in the project. The results support a vision towards an ILDE that enables teachers to choose among multiple learning design authoring tools, (co-)produce, explore, share, evaluate and implement learning designs in Virtual Learning Environments. The paper also outlines a roadmap to achieve this vision.

Keywords: learning design lifecycle, integrated environment, authoring tools, virtual learning environments, teacher communities.

1 Introduction

Learning design, as a field of research in TEL, is concerned with supporting teachers and course designers in analyzing, conceptualizing, and sharing the design of learning situations [1]. A rich set of learning design methodologies, languages, tools, and case studies are available today [2-3], but their transfer into classroom practice is still lagging behind expectations. There is a considerable body of research on why this is the case (see, e.g., [3-4]), including obstacles such as the lack of usable authoring tools, the lack of integration with real-world teaching processes, the lack of support during enactment of designs, and the lack of interoperability of learning design artifacts. Essentially, there is currently no integrated solution to support the full lifecycle of designing and deploying ICT-based learning activities. Furthermore, as with any other effort aiming to integrate ICT and other innovations in the classroom, adequate professional development of educators is recognized as being both difficult and crucial for success [5].

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 448–453, 2013.

The recently started "Meeting teachers' co-design needs by integrated learning environments" (METIS) project[1] aims to close this gap. It aims to mesh the current kaleidoscope of available learning design authoring tools, enhancing them to enable the sharing and enactment of learning designs (Ld), within an Integrated Learning Design Environment (ILDE). The project will also design professional development actions (e.g. workshops) to promote learning design practice, to be used in conjunction with the ILDE. To ensure broadest possible uptake, this initiative is methodologically based on strong involvement of end-user testbeds from different kinds of educational contexts. This paper presents the vision for the ILDE and the first design-based research studies involving end-users. These results will feed ILDE development and the production of workshop packages that will serve teachers outside of METIS in adopting ILDE as a tool for everyday practice.

Next section outlines the methodology adopted for gathering user feedback about the ILDE vision; the vision itself is presented in Section 3. Section 4 presents results from end-user consultations, and Section 5 concludes the paper outlining the roadmap towards the realization of the ILDE.

2 Methodology

The work presented in this paper corresponds to a first step of a design-based research methodology [6] where the vision of the METIS project for an ILDE is implemented and evaluated in real settings. This first step addresses the collection of early feedback regarding the need and relevance of this vision for the project user groups, represented by three educational institutions in the sectors of Vocational Training, Higher and Adult Education (VT: KEK, HE: The Open University, AE: Àgora)[2].

To facilitate the communication and inquiry about the relevance of the envisioned ILDE for the user groups, a generic scenario and a set of use cases capturing the vision were formulated. The collection of data combined the use of different data gathering instruments as summarized in Table 1. These instruments included questionnaires and on-line interviews focused on the generic scenario and uses cases, as well as a face-to-face session to work on more specific scenarios. The complete instruments and proposed scenarios are available in one of the project deliverables [7]. Next section presents the vision behind the ILDE along with a summary of the generic scenario and use cases.

Table 1. Data gathering instruments and type of data collected

Instrument	Type of data
Questionnaire 1, administered on-line, open questions [Q1]	Qualitative explanations about user groups' contextual information (3x Q1)
Questionnaire 2, administered on-line, closed and open questions [Q2]	Likert scales to rate the relevance of the use cases to their institution and open comments to explain the rationale behind the ratings
Interviews, via videoconference, related to questionnaire 2 [I]	Clarifications regarding use cases, qualitative data collected as comments associated to the ratings (3x1)
Face-to-face group work (with representatives of user groups and other project stakeholders) [GW]	Narratives with specific scenarios significant to the institutions of the user groups (1xGW)

[1] `http://www.metis-project.org` (Last visit: 8 Apr 2013)

[2] `http://www.eurotraining.gr/`, `http://www.open.ac.uk/`, `http://www.edaverneda.org/` (Last visit: 8 Apr 2013)

3 Vision for an Integrated Learning Design Environment

With the ultimate aim of fostering and sustaining the adoption of learning design, the vision for the ILDE is to support the complete learning design lifecycle: from authoring the designs to deploying them in VLEs for enactment and eventual redesign, all in the context of teacher communities. The support for this whole lifecycle is addressed from the perspective of the "user experience" in the flow of completing the various tasks involved in the lifecycle (i.e., the emphasis is not on achieving interoperability of designs between authoring tools).

To support teacher communities, the LdShake platform [8] is considered as one of the central elements of the ILDE (Fig. 1). It provides social network features, acts as a repository, controls the access to the designs, and presents – embedded in its Web space – Ld conceptualization and authoring tools. Depending on the nature of the authoring tool (document-based, Web-based, standalone desktop application), the integration approach with LdShake will vary (rich text templates, adapters, upload feature). Examples of tools that are to be integrated first (provided by project members) are Web Collage [9], OpenGLM [10], CADMOS [11], and conceptual design tools proposed by the OULDI project [2]. However, other tools are also under analysis [7] for their potential integration. The ILDE architecture also includes GLUE!-PS [12], a system to enable the deployment of designs in mainstream VLEs.

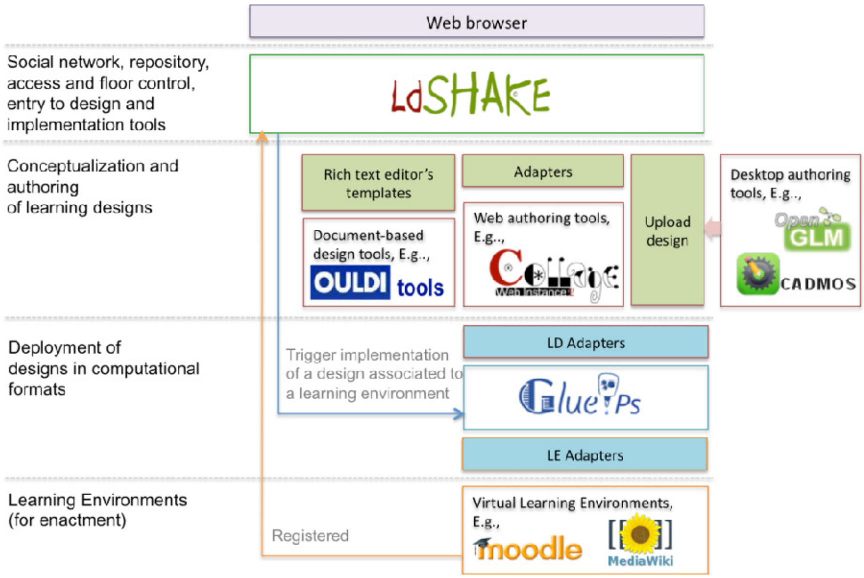


Fig. 1. ILDE architecture to support the whole learning design lifecycle

The vision for a generic scenario applying the ILDE includes the use cases listed in Table 2 which, if considered in the order presented, represent the typical flow of usage of the ILDE (even if multiple variations are possible – e.g., use case 8 can be used at different moments along the flow). The first end-user feedback about different manifestations of the use cases (see sub-cases in Table 2) is discussed next.

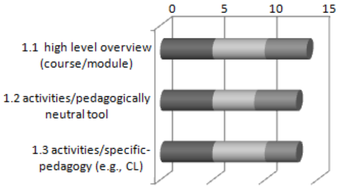
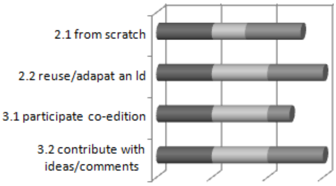
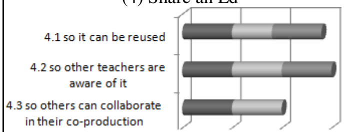
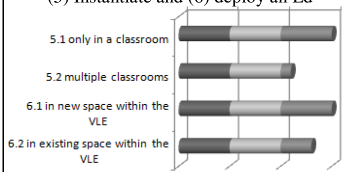
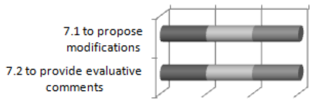
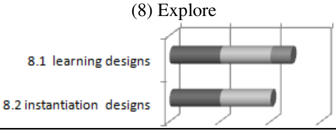
Table 2. Main uses cases considered in the vision for the ILDE

Use case	Description
(1) Choose a learning design tool	Depending on the purpose and needs of particular educational cases, teachers may require different learning design authoring tools. Besides, the iterative process involved in learning design may also lead to the same team / individual choosing different tools at different phases [1-3]. Teachers' goals may vary widely, including the design of (1.1) a high level overview of a course, (1.2) activities designed with a pedagogically-neutral authoring tool, (1.3) activities created following a pedagogically-specific design process such as collaborative learning, etc.
(2) Produce or (3) co-produce a learning design	When creating a learning design, two main desires from teachers are [3], [9-11]: (2.1) to produce a design from scratch, (2.2) to adapt and reuse an existing design. When participating in a co-design, there are two main possible situations: (3.1) participating in its actual co-edition, or (3.2) contributing with ideas (e.g. as comments) to the design.
(4) Share a learning design	Different teachers may have different intentions when sharing learning designs with other teachers [8], so that: (4.1) the design can be reused; (4.2) other teachers are aware of their designs; (4.3) other teachers can collaborate in their co-production.
(5) Instantiate and (6) deploy a learning design into VLEs	As a second step in the lifecycle, the designs need to be bound to the specific students (and other participants) and ICT tools that will support the activities in the designed learning experience [12]. The design may be applied (5.1) only to a specific cohort or classroom of students, or (5.2) it may be applied to multiple classrooms. Then, the design needs to be deployed in the target ICT learning environment (e.g., an institutional VLE like Moodle) either as (6.1) a new space (course) in the VLE, or as (6.2) additional activities to an existing space.
(7) Provide feedback and reflections	Teachers often observe aspects that need to be considered improved in a design. Formative or summative evaluation of designs (e.g. designs under analysis, or already used in practice with students) is essential for practitioners [1]. As a result, teachers may want to (7.1) modify a design or (7.2) provide evaluative comments about the enactment of a design for future consideration (related to cases 2-3).
(8) Explore designs, instantiations and feedback	In order to enable certain previous use cases, such as (3), and to complete the ultimate aim of others, such as (4) or (7), it is important to explicitly consider this use case including the possibility to: (8.1) explore designs, and (8.2) explore instantiations of designs (see case 5) available in the teaching community.

4 First Feedback from User Groups

The data collected (Table 3) indicate that the three user groups (from the sectors of AE, HE, VT) are interested in using a variety of Ld tools supporting different granularities (from modules/courses to activities) and design phases (from conceptualization to consolidation). Institutions whose designs have to be approved by authorities (VT) value especially the support to document high-level designs. The three user groups are highly interested in being able to (co-)create designs by reusing existing ones. This practice is preferred to creating new Lds from scratch, when possible. The three institutions recognized as an essential feature allowing diverse actors (also students –AE–, and managers –HE, VT–) to annotate the designs. Yet, the relevance of the co-production use case depends on the teacher collaboration culture of the institution/sector. While collaboration, sharing and reuse with multiple groups of students is critical in the studied cases of AE and HE, in the case of VT their applicability is very limited. Support for the whole lifecycle (including deployment and evaluation) will simplify processes in the three institutions. Besides, enabling the exploration of designs and comments (about the designs and their implementations with students) is valued as very useful, especially in the cases of AE and HE.

Table 3. Relevance/usefulness of the use cases for the user groups

Use cases, ratings [scale: 0 – not useful, …, 5 – very useful] Àgora (Adult E.) OU (Higher E.) KEK (Vocational T.)	**Qualitative data (selection)** [labels indicating instruments in Table 1 and user groups]
(1) Choose a learning design tool 	- Àgora's courses are composed of multiple small parts, having the global plan is important [Q2-I_Àgora] - We applied varied pedagogies [Q2-I_Àgora], though our main methodology (dialogic learning) [Q1_Àgora] involves collaborative learning [Q2-I-WG_Àgora] - High level designs are useful during the initial stages of design, and can also help explain to others how a design works, and /or how lower level designs fit together [Q1-Q2-I-WG_OU] - Documents with high level designs are submitted to the ministry [Q1-Q2-I-WG_KEK] - Guidance for considering specific pedagogical approaches would be useful for the design of micro-sessions (activities) [Q2-I-WG_KEK]
(2) Produce or (3) co-produce an Ld 	- Reusing is a common practice at Agora, we create from scratch only when starting courses on new topics [Q2-I_Àgora] - An aim for Agora is that more than 50% of its courses involve more than 1 teacher in their design [Q2-I-WG_Àgora] - It's important that students can contribute with feedback [Q2-I_Àgora] - Innovation requires building on existing knowledge, but there must be a way of assessing the designs [Q2-I_OU] - Teams create most designs [Q1-Q2-I-WG_OU] - There is a lack of culture of teacher collaboration [Q1-Q2-I_KEK] - Training managers need to make annotations [Q2-I- WG_KEK]
(4) Share an Ld 	- Reuse of designs is very relevant, some are common to several courses [Q2-I_Àgora] - Transparency is critical in Àgora [Q1-Q2-I_Àgora] - It's important that it does not imply extra work to teachers [Q2-I_Àgora] - We already use some Ld tools but they do not support co-editing and sharing [Q2-I_OU] - Sharing, only within KEK [Q2-I-_KEK]
(5) Instantiate and (6) deploy an Ld 	- The same activity is potentially applied to many groups [Q2-I_Àgora] - The ILDE could help us to optimize resources, simplify processes [Q2-I-WG_Àgora] - This is the typical procedure at the OU: a course typically run twice a year, for many years with only minor changes [Q1-Q2-I-WG_OU] - Designers need to consider VLE limitations beforehand [Q2-I_OU] - The designs are created for and applied to disadvantaged and socially excluded groups. Therefore, we often use the one-fits-to-all approach (a design for just a group) [Q1-Q2-I_KEK]
(7) Provide feedback and reflections 	- Very important given the (high) number of teachers in Agora, and that many of them change from year to year [Q1-Q2-I_Àgora] - There is a need of mechanisms to record how well a design worked in a particular context (will increase the quality of future designs) [Q2-I_OU] - It would be useful to have a set of evaluation criteria [Q2-I- WG_KEK]
(8) Explore 	- Designs can offer ideas, inspiration to others [Q2-I_Àgora] - It would be nice to comment the instantiations –enacted designs- (and be able to explore those comments) [Q2-I_Àgora] - There is no tradition in sharing Lds in continuing education and training. However, internal use and re-use of Lds will be very useful (if they can be filtered by subject domains) [Q2-I- WG_KEK]

5 Conclusions and Future Work

The first studies with end-users of different educational sectors support the vision for an ILDE covering the whole learning design lifecycle in the context of institutional teacher communities (from exploring designs, choosing an authoring tool, co-producing a design, sharing it, to implementing it into a VLE, or providing feedback and reflection) (details in [7]). The analysis also indicates that the user groups require support for different granularity levels and phases of the design process.

The roadmap towards the realization of the ILDE considers four cycles, in which iterative releases of the ILDE will be used in workshops with teachers of the different sectors. The first release will provide basic support for the analyzed use cases, following the architecture in Fig. 1. Advanced support for users' requirements and progressive tool integration will be considered in the following cycles. The roadmap will be fed with the results from the workshops, and an extended analysis of the use cases' relevance to the project associate partners. This extended analysis will indicate the degree of representativeness of the project for user groups in their sectors, and allow the project to provide an ILDE (and associated workshop packages) suitable for fostering and sustaining the adoption of learning design beyond the project.

Acknowledgments. This work has been partially funded by EACEA, METIS Project 531262-LLP-2012-ES-KA3-KA3MP. The authors want to acknowledge the contributions from other project members, who participated in the collection of data or reviewed earlier versions of this work (A. Brasher, M.A. Serrano, S. Retalis, Y. Mor, L. Samper, M. Boloudakis, G. Conole, F. Pozzi, Y. Frangou, P. Abenia, E. Gómez, A. Martínez, J. Blat, D. Persico, Y. Dimitriadis).

References

1. Conole, G.: Designing for Learning in an Open World. Springer, Heidelberg (2012)
2. Cross, S., Galley, R., Brasher, A., Weller, M.: OULDI-JISC project evaluation report: The impact of new curriculum design tools and approaches on institutional process and design cultures. OULDI Project (2012), `http://oro.open.ac.uk/34140/`
3. Lockyer, L., Bennett, S., Agostinho, S., Harper, B.: Handbook of Research on Learning Design and Learning Objects: Issues, Applications, and Technologies. IGI, Hershey (2009)
4. Neumann, S., Klebl, M., Griffiths, D., Hernández-Leo, D., De la Fuente-Valentin, L., Hummel, H., Brouns, F., Derntl, M., Oberhuemer, P.: Report of the results of an IMS learning design expert workshop. I.J. Emerging Tech. for Learning 5(1), 58–72 (2010)
5. Lawless, K.A., Pellegrino, J.W.: Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. Review of Educational Research 77(4), 575–614 (2007)
6. Amiel, T., Reeves, T.C.: Design-based research and educational technology: rethinking technology and the research agenda. Educational Techn. & Society 11(4), 29–40 (2008)
7. Hernández-Leo, D., Chacón, J., Prieto, L.P., Asensio, J.I., Derntl, M.: D2.1 Early feedback on ILDE requirements. METIS Project Deliverable (2013), `http://metis-project.org/`, `http://metis-project.org/`
8. Hernández-Leo, D., Romeo, L., Carralero, M.A., Chacón, J., Carrió, M., Moreno, P., Blat, J.: LdShake: Learning design solutions sharing and co-edition. Computers & Education 57(4), 2249–2260 (2011)
9. Villasclaras-Fernández, E.D., Hernández-Leo, D., Asensio-Pérez, J.I., Dimitriadis, Y.: Web Collage: an implementation of support for assessment design in CSCL macro-scripts. Computers & Education (in press)
10. Derntl, M., Neumann, S., Oberhuemer, P.: Propelling standards-based sharing and reuse in instructional modeling communities: The open graphical learning modeler (OpenGLM). In: 11th Int. Conf. Advanced Learning Technologies, pp. 431–435. IEEE Press, NY (2011)
11. Katsamani, M., Retalis, S.: Making learning designs in layers: The CADMOS approach. In: IADIS Conference on Computer Science and Information Systems, pp. 305–312 (2011)
12. Prieto, L.P., Asensio-Pérez, J.I., Dimitriadis, Y., Gómez-Sánchez, E., Muñoz-Cristóbal, J.A.: GLUE!-PS: A multi-language architecture and data model to deploy TEL designs to multiple learning environments. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 285–298. Springer, Heidelberg (2011)

A Modeling Language to Describe Reusable Learning Processes to Achieve Educational Objectives in Serious Games

Eike M. Hirdes and Jan Marco Leimeister

Kassel University, Information Systems, Kassel, Germany
{hirdes,leimeister}@uni-kassel.de

Abstract. Serious games combine motivational aspects of games with pedagogical approaches to enhance their players' learning performance. A major challenge in serious game development is the proper alignment of concepts coming from the two domains, game development and pedagogy. As a consequence, creating successful serious games is difficult for single individuals only well versed in one domain, but not knowing about the other. Existing approaches focus on formalizing learning processes within individual games, but are incapable of making the instructor's knowledge reusable in other serious games. To address this issue, we present a modeling language to capture pedagogical knowledge in reusable learning processes that enable game developers to create serious games that support imparting educational objectives. As a result, our language helps game developers to integrate educational approaches into serious games. The paper illustrates the model and structure of the language.

Keywords: serious games, modeling language, game development, educational objectives.

1 Motivation

Games can increase the effectiveness of learning processes [1]. Educational objectives are integrated within games, so that users can reach these playfully, while the nature of the game incites and motivates users [2]. These are so-called serious games (SG), digital games which engage the user and instruct him.

To enhance the players' learning performance and simultaneously achieve learning outcomes (educational objectives), pedagogical methods and approaches need to be integrated into the digital game's story [3]. During the development process of a SG the toughest challenge is the alignment of pedagogy and the game itself [4]. Two domains need to be taken into account, game development, which deals with motivation of the user, and pedagogy, which aims to achieve learning outcomes. As a consequence, creating successful SGs is difficult for single individuals only well versed in one domain, but not knowing about the other. But the transition of educational methods and techniques into SGs has not been done in a structured and standardized way [5]. The aim of our work is to facilitate and stimulate the transition of pedagogy to

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 454–459, 2013.

SGs. SGs developers should be able to rely on formalized learning processes for SGs based on activities linked to educational theories and models. Furthermore, these processes should not describe the actual 'run' of a learning process. Thus, they should be reusable without having specific learning and game content, which will be added when deploying these processes during development.

For this purpose a modeling language is required to create a standardized description of learning processes to achieve educational objectives within SGs. Our work aims to answer the following research questions: What are the requirements for a modeling language to describe reusable learning processes to achieve educational objectives in SGs? What are the structure and elements of the prior mentioned modeling language?

The structure of this paper is as follows: Section 2 presents the requirements for the modeling language, the structure and elements of which will be presented in section 3. Finally, section 4 presents conclusions, limitations and future work.

2 Requirements to Describe Learning Processes in SGs

The modeling language to describe reusable learning processes to achieve educational objectives in SGs should take into account the requirements (RQ) described in this section. The requirements where derived upon reasoning and the study of literature. Firstly, we describe **General Requirements**.

- **RQ1 Completeness:** The modeling language must be able to completely describe a learning process to achieve an educational objective, including all activities and the relationships between those. [6]
- **RQ2 Abstraction:** The modeling language must be able to describe learning processes as generalizations to achieve educational objectives. Thus, they should be free from specific learning and game content. The processes can be instantiated and adapted to the developers needs during SG development process. [6]
- **RQ3 Precondition:** Preconditions for a learning process should be specified to describe necessary characteristics of a SG, which need to be fulfilled to successfully deploy the learning process (e.g. single / multiplayer capability). [7]
- **RQ4 Reproducibility:** The modeling language should describe learning processes that repeated application in different game development processes is possible. [6]

Requirements for the Structure of Learning Processes define the structure and guidelines for the composition of learning processes consisting of activities:

- **RQ5 Modular composition:** Learning processes to reach educational objectives must be constructed from activities. [7]
- **RQ6 Sequencing:** The order in which the activities of a learning process have to be performed needs to be specified. This order ensures that all activities of an educational model implemented within a SG are followed in the correct sequence to achieve learning outcomes. [7]
- **RQ7 Selection:** Activities within a learning process should be able to be specified as optional. During instantiation of the learning process, these blocks can be omitted by game developers to better adapt the learning process to a SG. [7]

- **RQ8 Reusability:** All activities of a learning process should be extracted and reused in other learning processes. [8]
- **RQ9 Recurring patterns:** Activities within a learning process could be bundled and named. This should reduce complexity and allow the reuse of frequently needed sequences of activities. [5, 9]

Requirements to Support Learning are characteristics to illustrate learning in SGs:

- **RQ10 Educational objective:** The learning outcome must be specified as an educational objective. This supports the focused composition of activities on a target and a measurement for the efficiency and effectiveness of a learning process. [5, 7]
- **RQ11 Pedagogical flexibility:** It must be able to describe learning processes that are based on different educational models. These models could include, among others, problem-based learning, anchored instruction, cognitive apprenticeships, reciprocal teaching, goal-based scenarios and project-based learning. [5, 6, 10]
- **RQ12 Presentation:** To reach educational objectives, the learner should be provided with learning stimulations and learning content. Learning processes should allow presentation of these to the user. [11, 12]
- **RQ13 Assessment:** Learning processes should provide a way to evaluate the learning progress of the user. Thus, assessments needs to be implemented which can be used to reflect the current progress (RQ14) or adapt the learning process to the user's needs (RQ15). [9, 12]
- **RQ14 Feedback:** Learning processes for SGs should provide feedback to inform the learner about learning / game progress. It can be given immediately on events in the game or actions taken by the user or delayed and summative. [9, 12, 13]
- **RQ15 Adaptation:** The learning process must be able to describe personalization and adaptation aspects, so that the succeeding activities within the process can be adapted based on the preferences of the user or upon assessment results. [6, 9]
- **RQ16 Roles:** Activities in learning process should allow different roles, especially learner and instructor. In this way, certain activities of the learning process can only be fulfilled by certain roles (i.e. evaluation of a written assessment can only be done by an instructor). [6]

Requirements to Support Engagement are used to describe user involvement, engagement and playfulness in learning processes:

- **RQ17 Non-linearity:** A learning process should allow non-linearity. Based on the non-linear structure of games, a learner should be able to follow two or more parallel active tasks or activities. [14]
- **RQ18 Rewards:** When completing a learning process partially or wholly a learner should be able to receive a reward, i.e. points or achievements. [13]

3 A Modeling Language to Describe Reusable Learning Processes in SGs

The raised requirements serve as a basis to describe elements and develop a model of a modeling language to describe reusable learning processes for SGs. According to our prior work, we are using the terms serious game pattern (SGP), serious game bricks (SGB) and serious game composites (SGC) for our modeling language [5]. SGBs are activities, which can be sequenced to model the flow within a SG (RQ5). Activities can be bundled for reuse and these bundles are called SGC (RQ9). SGCs can consist of two or more SGBs and SGCs. Finally, SGPs are specialized SGCs – also consisting of SGBs and SGCs. But contrary to SGCs, SGPs are associated with an educational objective and an educational model (RQ10, 11). These SGPs represent learning processes and are free from actual learning and game content [5] (RQ2, 4). Fig. 1 represents a model of the modeling language we propose.

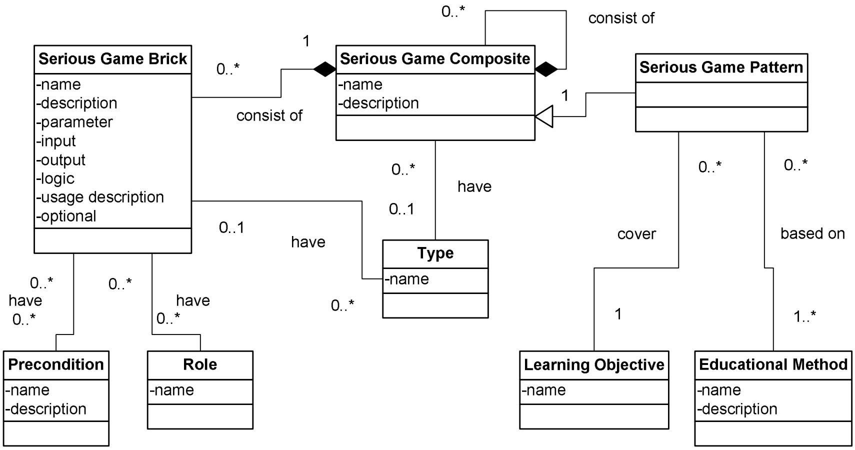


Fig. 1. Serious game learning process model

Table 1 offers a description to each class of our learning process model, their relation to each other, their associated attributes, as well as the requirements they are based on.

Table 1. Classes and attributes of the serious game learning process model

Class / *attribute*	**Description**
Serious Game Pattern (SGP)	SGPs describe learning processes and consolidate the actions a learner has to perform to attain learning objectives. (RQ1, 6)
Learning Objective	The learning objective describes the intended outcome of the learning process based on Bloom's revised cognitive domain [15]. (RQ10)
Educational Method	Provides the educational method the learning process is based on. (RQ11)
Serious Game Brick (SGB)	A SGB is a representation of an indivisible, basic activity of a SG. These SGBs can be arranged to describe gameplay and learning. (RQ5, 8)
logic	Describes how the parameter and/or input of a SGB are processed and describes the function of a SGB.
input	Input defines values obtained from other SGBs which can be processed. They are used as connection points for information flow connectors.

Table 1. *(Continued)*

output	If data is processes the output must be specified to pass the data to other SGBs. They are used as connection points for information connectors.
parameter	Parameters are interchangeable values to adapt the logic (i.e. answers of assessments). Parameters can be defined when deployed in an SG or changed through the input of prior SGBs. (RQ2, 4)
usage description	Usage description specifies rules for deploying SGBs. These are designed to support the developer to add learning / game content to SGBs by focusing on the attainment of the educational objective. (RQ2, 4)
optional	Defines whether a SGB is optional within a SGP or SGC, and thus could be omitted during deployment. (RQ7)
Precondition	Preconditions are capabilities of a SG which need to be fulfilled before a SGB can be deployed in a SG. (RQ3)
Role	Roles define users who have to fulfill the activities of a SGB, i.e. learner, evaluator or teacher. (RQ16)
Type	The type defines the purpose of a SGB or SGC. Type can be presentation, assessment, feedback or operator. (RQ12-14)
Serious Game Composite (SGC)	SGCs combine two or more SGBs or SGCs "to encapsulate several indivisible functions to one reusable complex function" [5]. (RQ9)
General attributes for classes	
name	Name of the class.
description	Description of the class.

Initial and final nodes are used to describe the start and end of a composition of activities within a SGP or SGC. Decision/merge and fork/join nodes are used to illustrate excluding and parallel activities (RQ15, 18). Finally, the language consists of two types of connectors [5]. The first kind of connectors is used to describe the logical order of activities (control flow connector) (RQ6). This allows the composition of SGBs to map educational methods to SGs. The second kind of connectors is used to handle the information flow between elements (information flow connectors), i.e. to pass results of an assessment to a feedback to allow adaption (RQ15). The elements of the modeling language are instantiated and combined to describe a SG dynamic, similarly to UML activity diagrams.

4 Conclusion and Outlook

This paper develops a description of learning processes by specifying and assembling activities linked to pedagogical objectives. The processes can be used during the development of SGs to attain educational objectives within the game. These should be able to fulfill all of the following: (1) describe processes to achieve educational objectives, (2) within SGs, and (3) are reusable and can be deployed repeatedly in various SGs. Based on literature as well as on observation of current practices, we raised requirements for learning processes. Based on these requirements we defined elements and developed a model for a modeling language to describe learning processes. The proposed processes can aid developers to implement educational objectives in a standardized way and, as a result, save time and money during development of SGs.

Our work shows the effect of principles between education and games and particularly, how learning objectives can be integrated as elements in games. Requirements and functional specifications are derived from the domain of technology enhanced learning. This paper also raises several issues. It must be examined, if the description of serious game patterns is sufficient to guide the game developers during development. The ease for reuse must be considered and evaluated through practical tests, to determine if the usage descriptions are sufficient. Further research is also needed to evaluate the use of patterns in various SGs.

Acknowledgment . This research was partially funded by the German Federal Ministry of Education and Research in the project ProduSE (www.projekt-produse.de), FKZ01FL10044.

References

1. Prensky, M.: Digital Game-Based Learning. McGraw-Hill, New York (2001)
2. Michael, D., Chen, S.: Serious games: Games that educate, train, and inform. Thomson Course Technology PTR (2005)
3. Greitzer, F.L., Kuchar, O.A., Huston, K.: Cognitive science implications for enhancing training effectiveness in a serious gaming context. Journal on Educational Resources in Computing 7, 10 (2007)
4. Harteveld, C., Guimarães, R., Mayer, I., Bidarra, R.: Balancing pedagogy, game and reality components within a unique serious game for training levee inspection. In: Hui, K.-c., Pan, Z., Chung, R.C.-k., Wang, C.C.L., Jin, X., Göbel, S., Li, E.C.-L. (eds.) EDUTAINMENT 2007. LNCS, vol. 4469, pp. 128–139. Springer, Heidelberg (2007)
5. Hirdes, E.M., Thillainathan, N., Leimeister, J.M.: Modeling Educational Objectives for Serious Games. In: Pedagogically Driven Serious Games Workshop 2012 (2012)
6. Koper, R., Manderveld, J.: Educational modelling language: modelling reusable, interoperable, rich and personalised units of learning. British Journal of Educational Technology 35, 537–551 (2004)
7. Janssen, J., Berlanga, A., Vogten, H., Koper, R.: Towards a learning path specification. International Journal of Continuing Engineering Education and Life Long Learning 18, 77–97 (2008)
8. Rawlings, A., Rosmalen, P., van Koper, R., Rodríguez-Artacho, M., Lefrere, P.: Survey of Educational Modelling Languages, EMLs (2002)
9. Marchiori, E.J., Del Blanco, Á., Torrente, J., Martinez-Ortiz, I., Fernández-Manjón, B.: A visual language for the creation of narrative educational games. Journal of Visual Languages & Computing 22, 443–452 (2011)
10. De Freitas, S., Jarvis, S.: A Framework for developing serious games to meet learner needs.In: The Interservice/Industry Training, Simulation & Education Conference. NTSA (2006)
11. Tang, S., Hanneghan, M., El-Rhalibi, A.: Pedagogy Elements, Components and Structures for Serious Games Authoring Environment. In: 5th International Game Design and Technology Workshop, Liverpool, UK (2007)
12. Gagné, R.M.: The Conditions of Learning and Theory of Instruction. CBS College Publishing, New York (1985)
13. Schell, J.: The Art of Game Design: A book of lenses (2008)
14. Garris, R., Ahlers, R., Driskell, J.E.: Games, Motivation, and Learning: A Research and Practice Model. Simulation & Gaming 33, 441–467 (2002)
15. Krathwohl, D.R.: A revision of Bloom's taxonomy: An overview. Theory into Practice 41, 212–218 (2002)

Motivational and Affective Aspects in Technology Enhanced Learning: Topics, Results and Research Route

Teresa Holocher-Ertl[1], Christine Kunzmann[2], Lars Müller[3], Verónica Rivera-Pelayo[3], and Andreas P. Schmidt[4]

[1] ZSI – Zentrum für Soziale Innovation, Austria
holocher@zsi.at
[2] Pontydysgu, UK
kontakt@christine-kunzmann.de
[3] FZI Research Center for Information Technology, Germany
{lmueller,rivera}@fzi.de
[4] Karlsruhe University of Applied Sciences, Germany
andreas_peter.schmidt@hs-karlsruhe.de

Abstract. Motivational and affective aspects have long been neglected in research and development of technology enhanced learning (TEL) solutions, but it is now increasingly recognized that they are key to acceptance and sustainable success. However, the consideration of these aspects still suffers from fragmented research activities that are in between established disciplines. We summarize the results from three editions of the EC-TEL workshop series MATEL, which has established a forum for interdisciplinary conversations and joint research activities. This includes an overview and systematization of current research and its findings as well as prioritized research challenges. The paper concludes with a research agenda that advances the inclusion of motivational and affective aspects into TEL from art to an engineering approach.

Keywords: Motivation, Affective, Emotions, Workplace Learning.

1 Introduction

Motivational and affective aspects are frequently neglected in TEL - although experiences from research and practice consistently show that they are one of the most important acceptance and success factors of learning solutions. This becomes even more important as we move towards more open, independent, and informal learning settings. Addressing these aspects requires a multi-disciplinary perspective, including pedagogy/andragogy, psychology, human resources management, computer-supported cooperative work, knowledge management, serious games engineering, affective computing and sensors engineering, among others. To provide a forum for such multi-disciplinary activities, to collect the scattered state of the art in the various fields, and to stimulate further research, the authors have initiated a successful workshop series at EC-TEL. This paper highlights findings from the first three editions of the workshop series (in 2010, 2011, and 2012). Each workshop consisted of research

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 460–465, 2013.

papers on results, short papers on on-going work, position papers for highlighting research directions, and tool demonstrations. Additionally, there were extensive interdisciplinary discussion sessions including "landscaping activities" of key concepts and identifying key challenges for further research (see for more extensive material under http://matel.professional-learning.eu). The accepted submissions were 8 for 2010 (22 participants), 5 for 2011 (17 participants) and 6 for 2012 (17 participants). The following sections highlight outcomes of the workshop series and present the identified research challenges, both for the motivational and the affective strand. We conclude by discussing the link between affective and motivational aspects.

2 Motivational Aspects

As a result of the discussions at the workshops, we could identify three main types of motivation relevant to TEL:

The **motivation to learn** addresses the question on what moves or hinders learners to learn, what makes them persevere in the face of difficulties and how the learners' motivation is influenced by teachers/tutors, practices as well as peer behaviour. The motivation to learn is understood as an individual's characteristic. This can be connected to concepts like learning outcomes, attitudes and self-efficacy [1, 2]. The project IntelLEO focused on a goal driven approach and developed a model on how to increase the motivation of self-regulated learners at the workplace. Holocher-Ertl [3] presented the evaluation results of TEL tools built according to this model which showed that self-regulated learning activities are important for self-motivation. Attention has to be given to potential conflicts between staff and management regarding the usage of the collected learning information, as well as to efforts spent with these activities. The project suggested putting further research in the investigation of "lightweight" tools for self-regulated learners and their role on motivation to learn.

The **motivation to share** knowledge is increasingly important where learning becomes more social. But challenges arise on how to make tacit knowledge of experts explicit and useful for other workers and how to motivate knowledge sharing behaviour. Studies focus on drivers and barriers [4] for knowledge sharing. The motivation to share knowledge was the aspect, which was most intensively addressed by contributions in the MATEL workshops. IntelLEO identified the importance of collaboration on the motivation to share knowledge and self-efficacy [3]. Cress [5] discussed the aspect of free-riding, which occurs when people read others' contributions but do not actively contribute because of manifold barriers. Further research in experimental settings investigated how far bonus systems, social norms, etc. influence people's contribution behaviour. Within MATURE, Kunzmann & Schmidt [6] have developed an analysis model for identifying motivational barriers in a concrete work context, with the perspectives (i) individual (values, interest, needs, and capabilities), (ii) interpersonal (cooperative, affective), and (iii) enabler (organization, infrastructure). Several concrete barriers were identified based on ethnographic fieldwork [7] and theoretical considerations, and the analysis model was embedded into a proposal for an iterative design process which has been used by Ravenscroft et al. [8]. Based on the model of Kunzmann & Schmidt [6], Cook et al. [9] extends this perspective further towards larger networks of users where direct involvement is no longer possible.

The **motivation to use a specific tool** is related to the general question of technology acceptance. Research in this area investigates factors that influence future users' acceptance of innovative tools and reflects on supportive research and user-involvement methodologies. One approach in the workshop was gamification, which makes use of game mechanics, such as giving scores for user activities and introducing leader boards for stimulating competitive behaviour [10], but it also has to acknowledge gender and cultural differences [11].

During each of the MATEL workshops, there was extensive space for discussion. Next to the motivational landscapes, the main outcome from these discussions was the identification of **research challenges** (see Fig. 1).

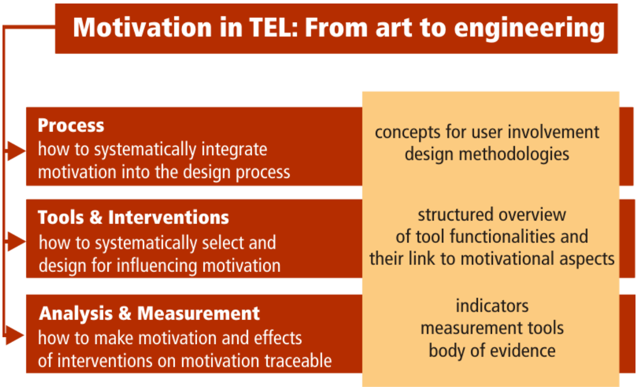


Fig. 1. Research challenges in TEL with respect to motivational aspects

The main challenge is advancing this topic *from art to engineering*. Motivational research is often dealt with in different contexts, some research is based on theories on motivation, others focus on empirical investigations, some address motivation via tools and organisational interventions, others via the design process. It is still a collection of detached case studies and ideas on how to solve problems in a certain domain.

We have to search for synergies between researchers and advance our joint knowledge on motivation to provide systematic approaches (based on a sound body of evidence) on three levels: (i) the level of *design processes*, where empirically grounded guidelines are needed on how to involve the future users into the design process in different contexts fostering aspects like autonomy, ownership and acceptance of innovative solutions; (ii) the level of *tools and interventions* where an inventory of functionality and other measures is needed, similar to design pattern approaches, and (iii) the level of *analysis & measurement* where indicators and measurement methods are needed to make motivation more traceable and to assess the effects of interventions.

On all three levels there is the aspect of *context*, which influences processes, tools and measurement. The challenge is finding ways to keep the complexity manageable while on the other hand not completely neglecting the different contexts where motivation is key to success. Towards that end we will need more specific investigations in comparative research between different approaches, preferably within a similar environment, which is often hard to achieve at the workplace.

3 Affective Aspects

In contrast to motivational aspects, affective aspects in TEL are still an emerging topic. There are first prototypes and studies (e.g., [12], [13]), but a common understanding of the role of emotional aspects is challenging and not achieved yet. There is a broad base of theory from psychology, but these theories are quite complicated and hence difficult to apply in a design process. The Circumplex Model of Affect by Russell [14] and the model by Krathwohl [15] are two theories that are often considered in TEL. Both provide an understanding of emotions and offer a starting point to design a technological support. The affective domain of education is a relatively little understood phenomenon, especially when considering technology support, and several tools are introduced in learning contexts. A discussion is presented in [16] along two conceptualizations of the affective domain of learning.

As part of the MATEL series, the main contributions were related to specific approaches to apply technology in a learning context, based in a certain psychological theory (or a part of it). These approaches presented three purposes for which emotions are taken into account in TEL:

- A first research direction investigates how emotions can be used for **adaptive systems** [13], which adapt themselves and react according to the emotions of the user.
- **Raising awareness and making emotions available for reflection** (as content of the reflection itself) [12], [17] is another trend of employing affective aspects in learning. Being aware of emotions and reflecting on them can be supported both at individual as well as collective level.
- Affective aspects in TEL are considered in approaches **to influence and regulate emotions of users** [18].

The goal of tackling affective aspects in TEL was also a discussion topic during the MATEL workshops. This goal is still undefined and it is unclear if the aim should be to feel and induce always positive emotions, similar to motivation, where the main goal is to achieve a high motivation. In the case of reflective learning, not only positive emotions are considered, but negative emotions play an important role as well. The discussed TEL solutions are facing similar problems, because they tend to follow one single theory coming from psychology, and this complicates the definition of a common understanding. The main challenges are (i) the lack of a common language, (ii) the dependency on the users' context and (iii) the specific barriers and problematic assumptions for the success of applications.

The discussion of a common understanding of affective aspects was an on-going theme during all MATEL workshops. A *common language is lacking* to connect the different involved disciplines, as designers, psychologists and sensor experts can use completely different terms. The communication between an application and the user suffers from a similar problem, because they depend on the *visualization* of the data, e.g. an icon that is selected to express mood can be used to communicate the current mood to others or store it for later reflection. Textual or graphical descriptions create different notions depending on the reader's background. There is a wide variety of user interfaces that support the reporting of subjective impressions. In the field of

HCI, interfaces and representations for emotions have been extensively researched, but we still lack guidance on which methods or visualizations are more suitable for which contexts and purposes, especially in learning. There is not a *universal preference to express and communicate* emotions from the user's point of view and some users may prefer to use smileys, while others prefer a term from a taxonomy. This fact adds barriers and challenges to developing tools that may offer support to all of them.

Finally, emotions are dependent on the current *personal context* e.g. the reaction of a person to a certain event can be different depending on the time of the day or the place where she is. This dependency originates that the meaning of the representations of emotions varies depending on the person's context.

Major *concerns or barriers* regarding affective aspects have been raised and discussed in the MATEL workshops. For instance, while humans are trained in detecting affective state of others, automated methods for detecting affective state will never be as good at it. This means that we should not seek to replace human perception, but rather to augment it. *Problematic assumptions* which may not be taking into account in research about emotional aspects of TEL are also discussed in [16].

4 Conclusions

The MATEL workshops between 2010 and 2012 have identified concrete research challenges to guide future research in the field. It has become obvious that emotions and motivation are linked to each other closely, although their relationship is not an easy one. Emotions are on a low level of abstraction and close to what can be observed; but rather short-lived and quick to change. Effects of emotions on learning processes are very complex as short-term effects often differ from long-term effects (such as the effects of stress, or anxiety), but the field of affective aspects still lacks a more stable notion of affective factors. Motivational aspects on the other hand are on a higher level of abstraction, which is more stable and makes it easier to link it to impact on learning processes. But motivation remains hard to observe, detect, and to influence. Motivation is influenced by emotions, aggregating series of emotions into a more stable motivational state. Due to its closeness to learning, research on motivational aspects has reached a higher maturity, while affective aspects have not yet a coherent research area. Overall, in both areas, the workshops have clearly highlighted the importance of contextual factors, ranging from individual characteristics, via learning context to tool context. To advance the field from an art to an engineering approach of designing motivational and affective aware systems, investigation of more concrete examples have been proposed to gain a richer body of evidence. Suitable engineering methodologies will probably be based on design-based research as a method and agile project management methodologies that refer to a shared conceptual model at the centre such as [19].

Acknowledgements. This work is supported by the European Commission under the FP7 projects IntelLEO (no. 231590), LAYERS (no. 318209), MATURE (no. 216356), and MIRROR (no. 257617). We thank previous members of the MATEL team for their support.

References

1. Montalvo, F.T., González Torres, M.C.: Self-Regulated Learning: Current and Future Directions. Electronic Journal of Research in Educational Psychology 2, 1–34 (2004)
2. Zimmerman, B.: Self-Efficacy: An Essential Motive to Learn. Contemporary Educational Psychology 25, 82–91 (2000)
3. Holocher-Ertl, T., Fabian, C.M.: The influence of collaboration and self-regulated learning services on the motivation to learn at the workplace. In: Proc. MATEL 2011-2012 (2012)
4. Riege, A.: Three-dozen knowledge-sharing barriers managers must consider. Journal of Knowledge Management 9, 18–35 (2005)
5. Cress, U.: Why people withhold information: Information exchange as social dilemma. In: Proc. MATEL 2010 (2010)
6. Kunzmann, C., Schmidt, A., Braun, V., et al.: Integrating Motivational Aspects into the Design of Informal Learning Support in Organizations. In: i-KNOW 2009, Graz, Austria, pp. 259–267. Springer (2009)
7. Kunzmann, C., Schmidt, A.: Ethnographically Informed Studies as a Methodology for Motivation Aware Design Processes. In: Proc. MATEL-2011-2012, ECTEL 2011, Palermo, Italy (2011)
8. Ravenscroft, A., Schmidt, A., Cook, J.: Designing for Motivation in TEL: Relevance, Meaning and Value in Context. In: MATEL 2010, ECTEL 2010, Barcelona, Spain (2010)
9. Cook, J., Schmidt, A., Kunzmann, C., Braun, S.: The Challenge of Integrating Motivational and Affective Aspects into the Design of Networks of Practice. In: Proc. MATEL-2011-2012 (2013)
10. Parra, G., Klerkx, J., Duval, E.: TiNYARM: This is Not Yet Another Reference Manager. MATEL 2012. In: Proc. MATEL-2011-2012 (2013)
11. Schell, J.: The Art of Game Design: A book of lenses. CRC Press (2008)
12. Fessl, A., Rivera-Pelayo, V., Müller, L., et al.: Motivation and User Acceptance of Using Physiological Data to Support Individual Reflection. In: Proc. MATEL-2011-2012 (2013)
13. Leony, D., Pardo, A., Parada, G.H.A., Kloos, C.D.: A Widget to Recommend Learning Resources Based on the Learner Affective State. In: Proc. MATEL-2011-2012 (2013)
14. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology 39, 1161–1178 (1980)
15. Krathwohl, D.R., Bloom, B.S., Masia, B.B.: Taxonomy of Educational Objectives Handbook II: Affective Domain. Evaluation 4, 164–170 (1965)
16. Voigt, C., MacFarlane, K.: The affective domain and social networking: definitorial issues and misleading assumptions. In: MATEL 2010, ECTEL 2010, Barcelona, Spain (2010)
17. Fessl, A., Rivera-Pelayo, V., Pammer, V., Braun, S.: Mood Tracking in Virtual Meetings. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 377–382. Springer, Heidelberg (2012)
18. Schaaff, K., Müller, L., Kirst, M., Heuer, S.: xAffect - A Modular Framework for Online Affect Recognition and Biofeedback Applications. In: Proc. MATEL-2011-2012 (2013)
19. Ravenscroft, A., Schmidt, A., Cook, J., Bradley, C.: Designing social media for informal learning and knowledge maturing in the digital workplace. Journal of Computer Assisted Learning 28, 235–249 (2012)

A Case Study of Interactive Tabletops in Education: Attitudes, Issues of Orientation and Asymmetric Collaboration

Andri Ioannou, Maria Christofi, and Christina Vasiliou

Cyprus University of Technology, Department of Multimedia and Graphic Arts, Limassol, Cyprus
andri.i.ioannou@cut.ac.cy,
{mc.christofi,vasi.christina}@gmail.com

Abstract. This paper is concerned with the exploration of an educational tabletop application designed to facilitate collaboration amongst young learners while they learn about the "Plants of Cyprus". The application was used by 28 third-graders during a scheduled visit at the Cyprus Center of Environmental Research and Education. We report empirical findings concerning the participants' interactions around the table as well as their attitudes regarding the activity. Findings demonstrated that the students collaborated intensively in completing the task and they were overwhelmingly positive about the experience. The paper discusses issues of orientation of the on-display learning artifacts, which encouraged learners to move at a new location around the table to "correct" the orientation. Also, the study raises concerns regarding asymmetrical forms of collaboration, where peers dominated the activity despite the equal access on the tabletop surface.

Keywords: Interactive Tabletops, Collaborative Learning, Interactions, Attitudes, Orientation, Asymmetric Collaboration.

1 Introduction

Multi-touch interactive tabletops have recently attracted the attention of the Human Computer Interaction and Educational Technology communities. A multi-touch interactive tabletop can handle multiple simultaneous touch inputs and can support collaboration by allowing different patterns of turn taking, negotiation and interaction [2][1]. As discussed by [4], multi-touch tabletops afford cooperative gestures which can enhance users' sense of teamwork. In this work, an educational tabletop application was designed to facilitate collaboration amongst young learners while they learn about the "Plants of Cyprus". The study sought to (1) explore the kinds of interactions evident around the tabletop and (2) examine students' attitudes toward the activity.

2 The "Plants of Cyprus" Application

The "Plants of Cyprus" aims to facilitate collaboration amongst young learners while they learn the different types of the plants growing in Cyprus and their uses in

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 466–471, 2013.

cooking, weaving, pharmacy and basketry. Two educators at the Cyprus Center of Environmental Research and Education were actively involved in the design process and pilot-testing of the application. The "Plants of Cyprus" runs on a TouchMagix table -- a multi-touch technology that supports multiple simultaneous users. The application is designed for four users (one at each side of the table) and is completed in two stages.

In Stage 1 (declarative knowledge), learners interact with the application to retrieve information through a series of videos concerning the various plants growing in Cyprus and their uses (see Fig. 1). In this stage, learners are advised to collect as much information as possible, before they proceed to Stage 2 of the application; Stage 1 is not accessible once learners proceed to Stage 2.

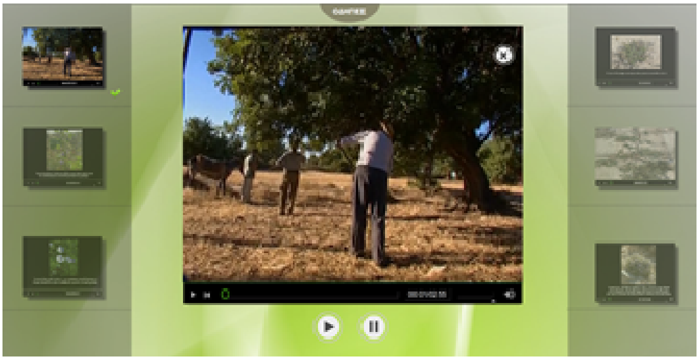


Fig. 1. "The Plants of Cyprus": Stage 1 – declarative knowledge

In Stage 2 (assessment), learners are asked to collaborate on a matching activity to assess their level of "recall" of declarative knowledge (see Fig. 2). The pictures of eight plants are presented and each one has to be linked to a category/use (cooking, weaving, pharmacy and basketry). Some indicative information about the plant is displayed to help learners recall the information from Stage 1. The application is collaboration-enforcing in the sense that learners have to discuss the information they recall from Stage 1 in order to match the plants to the correct category and complete the task successfully; students are allowed to try till they find the right solution.

Furthermore, the application is designed to promote symmetrical collaboration where all four learners have equal access to the tabletop and equal opportunity to match plants to categories. This is achieved through the design of a workspace for every user as shown in Fig. 2, as well as the restriction imposed by the application that each learner has to match exactly two plants. That is, a learner chooses two plants to link to the appropriate category; these plants cannot be reused by another learner.

3 Method

A convenience sample of 28, third-graders (8-9 years old) participated in this study during a scheduled visit at the Cyprus Center of Environmental Research and Education with their teachers (2 teachers).

Students worked in groups of four, formed by their teachers (i.e., 7 groups total). Groups, one after the other, were taken to a quiet place at the Center where they engaged with the "Plants of Cyprus". There were no time restrictions in completing the activity; all groups spent 25-30 minutes on task.

Students' interactions were videotaped for subsequent analysis. Also, a questionnaire was administered to the students at the end of the activity to assess their attitudes regarding the experience. The questionnaire included six items with a Likert-type response scale from 1: strongly disagree to 5: strongly agree in the form of smiley faces, appropriate for the age of the participants. Also, the questionnaire included an open-ended question regarding students' experience.

Fig. 2. "The Plants of Cyprus": Stage 2 - assessment

4 Analysis and Results

Video analysis was conducted first. Two researchers (coders), with professional backgrounds in educational technology considered the video in its entirety (approximately 3 hours) in an effort to categorize the types of discourse and actions present. The coders worked closely together to create coding categories on the basis of the data (i.e., a bottom-up approach). Table 1 presents the coding scheme of the study, including four categories of verbal and non-verbal behavior. Using this coding scheme, the researchers coded all video data. The unit of analysis was the "unit of meaning". The percent agreement between the coders was 85% for both segmentation (into units of analysis) and categorization. Finally, codes in each coding category were counted as presented in Table 2.

Table 1. Coding Scheme

Coding Category	Description	Example
Task/tool related talk	Information sharing, providing advice, seeking confirmation, general talk about the task, asking/commenting about the technology.	M3: "Why can't I drag this plant?" M2: "Because I have it…take a different one" (tool related, group 5) M4: "I know what a carob is. It is food…you can eat it. I can do this matching." (task related, group 2) M2: "Oh you took "savory"… The video said it is used in some recipes. Do you know how to match it?" M3: "Yes, it is food." (task related, group 3)
Questioning/answering	Specific content-related questions and answers amongst the participants.	M1: "Broom…How do we make a broom?" M2: "Someone in the video explained it. It has to do with basketry I think." (group 2)
Dominating talk/move	Asking to lead or physically blocking and controlling others' actions.	M1 is trying to match a plant to a category but M2 pushes M1's hand away from the surface to do it himself, while saying "No no no … let me do it" (group 7)
Body relocation	Moving body to face learning artifact at a proximally "normal" orientation.	Students move physically at a new location around the table to orient themselves towards the videos, images and text to ease reading/viewing.

Table 2. Counts Within Each Coding Category Across Groups

	G1	G2	G3	G4	G5	G6	G7	Total
Task/tool related talk	14	14	21	15	13	14	9	100
Questioning/answering	3	4	3	3	2	3	2	20
Dominating talk/move	9	12	13	8	10	11	4	67
Body relocation	1	1	1	1	1	1	1	7
Total	27	31	38	27	26	29	16	194

As the frequencies of Table 2 illustrate, the participants collaborated intensively overall, especially in Stage 2 of the activity. In particular, the participants shared information about the plants they recalled from Stage 1 or provided advice to their peers to help them find the correct matching. For example, in Stage 2, one participant advised another member: "Read the description of the plant and if it says that you can eat this plant you will match it to the cooking category" (task/tool related talk). Also, some talk concerned confirmation seeking (e.g. "I think this plant belongs here. I am matching it") or asking/commenting about the technology (e.g., "I cannot drag it, how do you do it?"). Specific content-related questions and answers were less frequent, although it was present in all groups. Interestingly, questions never went unnoticed by the cooperating peers, for example M3:"What does weaving means?" M4:"It is our cloths" (group 1).

Despite the intensive collaboration, in each group a peer dominated the activity by asking others to let him/her complete the tasks, for example, "Let me do it…I know the answer" or by blocking and pushing their hands away from the surface to touch him/her-self (dominating talk/move). Furthermore, in Stage 1 and throughout viewing the videos for information retrieval, students in all groups move physically at a new location around the table where they viewed the learning artifact at a proximally "normal" orientation (see Fig 3 left-side); then in Stage 2 students return to their working spaces for the matching activity (see Fig 3 right-side).

Fig. 3. Collaboration on the "Plants of Cyprus"

Following the video analysis, the data from the attitudes questionnaire was analyzed. Student feedback was overwhelmingly positive, with a mean of 4.42 (SD=.77) across the six items of the questionnaire, as shown in Table 3. Also, in the open-ended question of the questionnaire, many students expressed their enthusiasm about the tabletop (e.g., "I would like to have one at home!", "I would like to play again"). Overall, the results from the questionnaire informed the results of the video analysis showing that not only the completion of the task was a group effort, but it was also an enjoyable one.

Item	Mean (SD)
1. I enjoyed the activity around the tabletop.	4.79 (0.50)
2. I would like to use a tabletop for school activities.	4.44 (0.85)
3. The tabletop encouraged my participation in the task.	4.12 (0.93)
4. The tabletop encouraged my collaboration with other group members.	4.15 (0.90)
5. My group worked well for this activity	4.38 (0.90)
6. I learned new things from this activity	4.65 (0.56)

5 Discussion

Findings from this study demonstrated that the students collaborated intensively in completing the matching task of the "Plants of Cyprus" around the table. Students' interactions were rich in information sharing, offering advice, confirmation seeking, questions and answers and general talk about the task and the technology as evident in Tables 1 and 2. Our investigation confirms previous research findings discussing the affordances of interactive tabletops to support collaboration (e.g., [1][2][4]).Also, student feedback was overwhelmingly positive, consistent with previous works suggesting that tabletops can enhance users' sense of teamwork [4] and can improve the (learning) experience and motivation to engage in the task [1].

Furthermore, the study showed that the orientation of the learning artifacts on the tabletop encouraged learners to physically move at a new location around the table to "correct" the orientation (i.e., oriented themselves towards the artifact to ease reading/viewing). This moving occurred in all groups during Stage 1; then students moved to their working space for Stage 2. This result confirms previous work with non-computer based, table-centered collaborative tasks showing that users consider straight-on orientated text as more readable (e.g., [3][5]). This finding suggests that learners around a tabletop should be free to move around to "correct" the orientation of the learning artifacts; being seated or restricted within the physical space can limit their ability to read/view.

Last, the study revealed asymmetric forms of collaboration in all participating groups with one group member dominating the activity, despite the deliberate design of the application to promote equal access to the tabletop and equal opportunity to match plants to categories. This finding divergences, in part, from the premise that multitouch interactive tables can support collaboration by allowing different patterns of turn taking and interaction [1][2]. This result might be suggesting that some form of facilitation or guidance is necessary for young learners to engage in a well-balanced collaboration free of dominant talk and blocking moves by peers.

References

1. Buisine, S., Besacier, G., Aoussat, A., Vernier, F.: How do interactive tabletop systems influence collaboration? Computers in Human Behavior (in press)
2. Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., et al.: Around the table: Are multipletouch surfaces better than single-touch for children's collaborative interactions? In: Proc. International Conference on Computer Supported Collaborative Learning, ISLS, pp. 335–344 (2009)
3. Kruger, R., Carpendale, S., Scott, S., Greenberg, S.: How people use orientation on tables: Comprehension, coordination and communication. In: Proc. of the 2003 int. ACM SIGGROUP Conf. Sanibel Island, Florida, USA, pp. 369–378. ACM Press (2003)
4. Morris, M.R., Huang, A., Paepcke, A., Winograd, T.: Cooperative gestures: Multi-user gestural interactions for co-located groupware. In: Proc. CHI 2006, pp. 1201–1210. ACM (2006)
5. Tang, J.: Findings from observational studies of collaborative work. International Journal of ManMachine Studies 34(2), 143–160 (1991)

What Happened to the Crossdisciplinarity of Technology-Enhanced Learning in 2004?

Marco Kalz

Centre for Learning Sciences and Technologies (CELSTEC),
Open University of the Netherlands, P.O. Box 2960, 6401 DL Heerlen, The Netherlands
marco.kalz@ou.nl

Abstract. In a recent study the crossdisciplinarity of the field of Technology-Enhanced Learning was analysed with science-overlay-maps and diversity measures. Results reveal that the crossdisciplinarity of the field has constantly increased over the last 10 years. Only in 2004, a significant decrease of inter-disciplinary research could be identified. In this paper we take a closer look at the publications of this year and test our hypotheses for the decrease of cross-disciplinarity.

1 Introduction

Technology-enhanced learning (TEL) is defined as an interdisciplinary field of research to which a number of disciplines contribute, namely: cognitive science, educational psychology, computer science, anthropology, sociology, information sciences, neurosciences, education, design studies, instructional design and others. According to Sawyer the field has been established in the late 1980ies based on the recognition that new scientific methods are needed that go beyond their own research field [1]. There are currently only a few qualitative studies about interdisciplinarity of TEL showing a mixture of methods and academics backgrounds of researchers in the field [2, 3]. While these studies report relevant findings most studies have the limitations that their data basis is small and that the approach chosen is not scalable.

In this paper we report results from a recent study that uses science-overlay maps and diversity measures to analyse the crossdisciplinarity of the TEL field [4]. In this study a steady increase of crossdisciplinarity could be identified for the field of Technology-Enhanced Learning. Only in 2004 there was a drop of crossdisciplinarity. We use science overlay maps and diversity measures to analyse the reasons for this decrease. In the next part we summarize the method chosen, then we present results of the analysis and discuss our findings.

2 Method

2.1 Science-Overlay Maps and Diversity Measures

We follow in our study an approach proposed by Rafols, Porter & Lydesdorff [5]. The basis for the analysis is a scientometric analysis of datasets from the Web of Science

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 472–477, 2013.

(WoS) by Thompson Reuters. One of the most developed approaches that takes into account the dynamic structure of journals belonging to scientific disciplines and the citations inside and outside the discipline has been proposed by Leydesdorff [6]. Another important component of the analysis is the integration score proposed by Porter & Rafols [7]. This integration score is based on the idea that the level of interdisciplinarity can be assessed by an analysis of three different aspects:

- the variety of the field (number of disciplines cited),
- the balance of the field (distribution of citations between fields) and the
- the disparity of disciplines cited (how similar are these disciplines).

This combination of perspectives to measure interdisciplinary research is aligned with earlier approaches to measure diversity by Rao [8] and Stirling [9]. Hence, the presented integration score is a special form of the Rao-Stirling Index of Diversity. This index can be understood as “a Simpson diversity in which the products of proportions of categories are weighted by distance/similarity” [10]. The basis for this method is a global map of science constructed with the help of the subject categories (SC) in the Web of Science. Leydesdorff & Rafols [11] have constructed a matrix and global map of science with the analysis of citing Subject Categories and cited subject categories in the Science Citation Index (SCI) and the Social Science Citation Index (SSCI) (in total 221 SCs) resulting in 18 macrodisciplines and their relations after a factor analysis. Leydesdorff, Carley and Rafols report that the resulting maps have been proven to be stable compared to other approaches to build a global map of science [12]. This global map of science is now used and overlayed with a local map resulting from a specific search approach in the WoS databases. This method is thus used to locate a specific organization, individual or topic on the global map of science. To come to such a local map for an overlay, a citation analysis is conducted.

2.2 Procedures

In a first step we have conducted a search in the WoS version 5.6 in July 2012 for the period between 2002 and 2011 in the SSCI (Social Science Citation Index) and SCIE (Science Citation Index Expanded) databases with the following keywords: “learning sciences”, “technology-enhanced learning”, “computer-based training”, “elearning”, “e-learning”, “mobile learning”, “electronic learning environments” and “educational technology”. This query resulted in 4255 records. We have narrowed the results down to journal articles and proceedings papers and we have only included articles in English. The resulting list consists of 3490 records. This list has been manually checked if the articles belong to the field. 14 results have been manually deleted from the original list so that finally 3476 records have been used for further analysis. These records have been further analyzed according to the procedure described in Rafols, Porter, & Leydesdorff [5]. The list of records has been filtered according to the Web of Science Categories with a minimum threshold of 1 occurrence and exported into a local file. This file has been converted with the small applications described in the paper above. The resulting matrices have been visualized with the Pajek application for the analysis of large networks based on a 19 cluster solution [13]. In addition we have constructed an overlay map for each year of the dataset. To assess the development of the cross-disciplinarity of the field over time we have calculated a Rao-Stirling-Index of Diversity for each year of the dataset. For the analysis of the year 2004 we have in addition constructed a table that shows the impact of individual disciplines to the diversity score of this specific year.

3 Results

To show the development of the crossdisciplinarity over time we have visualized the results of the individual science-overlay maps for years of the analysis in a video[1]. This animation of the science-overlay maps shows that the field of Technology-Enhanced Learning cannot be easily reduced to disciplines as proposed by other authors but is much more diverse. To be able to measure the level of interdisciplinarity we present in figure 1 a plot of the number of publications and the Rao-Stirling Diversity Index for each year of the analysis.

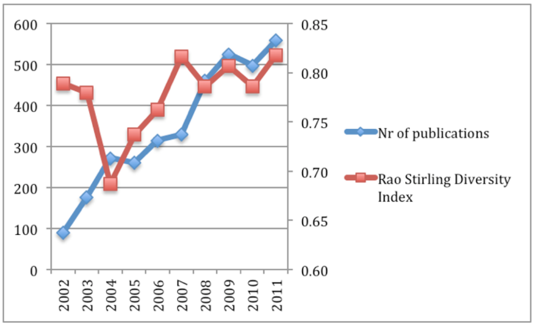


Fig. 1. Nr. of publications/Rao-Stirling Diversity Index

This figure shows that the crossdisciplinarity of the field has increased in the last 10 years. Only in the year 2004 a significant drop of the diversity index used in the analysis could be identified. Figure 2 shows the science-overlay map for 2004.

The science-overlay map for the year 2004 and the comparison to other years of the analysis revealed that this was the only year in which the output from computer science were higher than the educational sciences and other fields. Thus our hypothesis was that the decrease in interdisciplinarity was caused by the high amount of contributions from the field of computer science. To analyze the impact of contributions from individual domains we have conducted a more in depth-analysis for the dataset for the year 2004. We have stepwise omitted results from individual fields to control the effect on the Rao-Stirling-Diversity for this year. The results from this analysis are presented in table 1.

[1] `https://vimeo.com/46020529`

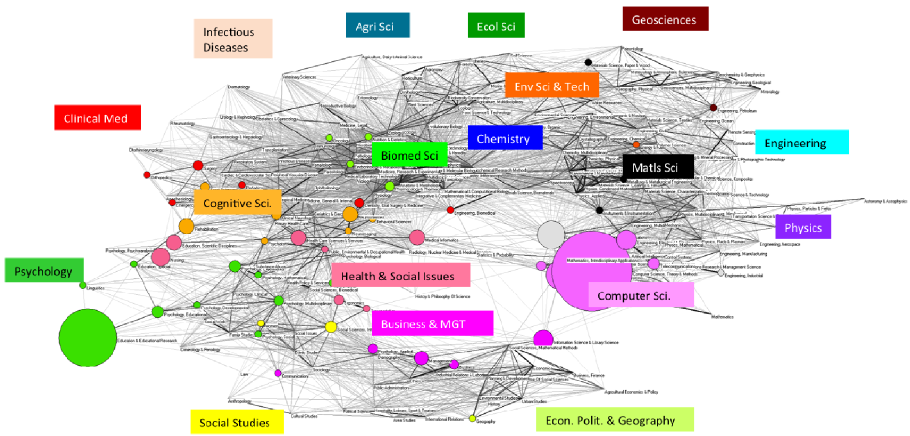


Fig. 2. Disciplinary composition of Technology-Enhanced Learning based on 271 citable items (journal papers & proceedings papers) published in 2004

Table 1. Impact of deletion of science categories on Rao-Stirling Diversity Index

Dataset	Nr. of publications	Rao-Stirling-Index
Full	368	0.69
Full - Educational Sciences	297	0.69
Full - Computer Science	179	0.86
Full - Health Sciences	343	0.67
Full - Business & Management	360	0.68
Full - Psychology	355	0.67
Full - Engineering	357	0.68

Table 1 presents the impact of deleting fields from the dataset. There is no impact of deleting the records from educational sciences science categories (Education: Educational Research & Education: Scientific Disciplines) although this set accounts for 20% of the full set of records. Deleting all publications from computer science from the dataset (with science categories CS: Theory & Methods, CS: Software Engineering, CS: Information Systems, CS: Interdisciplinary Applications, CS: Artificial Intelligence, CS: Hardware Architecture) has a huge impact. This subset accounts for approx. 50% of the full dataset and raises the Rao-Stirling-Diversity Index for the year 2004 from 0.69 to 0.86. The effect of subtracting publications from other fields like Health Sciences, Business & Management, Engineering or Psychology is very small. This can be also explained through the small set of publications for these fields for 2004. In the next paragraph we discuss our findings and we draw conclusions from the results of this study.

4 Discussion

The results of the study show that Technology-Enhanced Learning is a field to which a diverse set of disciplines contribute. The analysis of science-overlay maps produced in this study shows that the diversity of the field goes beyond the image that is paradigmatically reproduced in other publications. As an example the STELLAR Network of Excellence recently described the TEL community as consisting of a “more technical-centered silo and on the other hand a more people-centred silo” [14]. The reproduction of this old paradigm can be rejected based on the analysis. The diversity of the TEL-field is larger and cannot be described by such a simple model. The Rao-Stirling-Diversity Index score has shown that the crossdisciplinarity of the field has constantly increased as well as the overall scientific output of the field.

The decrease of crossdisciplinarity could be explained due to the high amount of publications from the field of computer science. 2004 was the only year of the analysis where the output from computer sciences and specifically the sub-category “Computer Science Theory and Methods” about TEL exceeded even the educational sciences. A closer look at the publications shows that most of these stem from a conference on adaptive hypermedia in the “Lecture Notes in Computer Science” Series. The references cited in these paper are clearly rooted in computer sciences and thus the citations out of the discipline are rather small. Thus this output has decreased the level of crossdisciplinarity in this specific year.

5 Conclusions

The study has shown that crossdisciplinarity of Technology-Enhanced Learning cannot be treated as a static entity but is a dynamic factor that is influenced by publications trends and citation practices of actors contributing to the field. We believe that the method applied in the analysis is an accessible and scalable way to measure interdisciplinary research conducted in the field. There is a raised interest of policy-makers to make interdisciplinarity as a requirement for funding decisions. The measures applied in this study have the potential to inform policymaking and allow a more objective assessment of scientific actors (people, institutions, disciplines) in the TEL field and related domains. In a recent study in the domain of Innovation Studies and Business & Management Rafols, Leydesdorff, O’Hare, Nightingale, & Stirling [15] show that the strong focus on journal rankings and impact measurement can in fact suppress interdisciplinary research. Therefore we see it as an important issue that policy makers are able to operationalize and measure the actual crossdisciplinarity of work funded. The method introduced here can build a basis for such an assessment.

References

1. Sawyer, R.K.: Optimising learning: Implications of learning sciences research. OECD, Paris (2008)
2. Conole, G., Scanlon, E., Mundin, P.: Technology enhanced learning as a site for interdisciplinary research. Report for the TLRP TEL programme (2010)

3. Plowman, L., Misailidou, C., Laurillard, D.: Developing Research Capacity for Technology-Enhanced Learning, London, UK (2007)
4. Kalz, M., Specht, M.: Assessing the crossdisciplinarity of Technology-Enhanced Learning with Science Overlay Maps and Diversity Measures. British Journal of Educational Technology (in press)
5. Rafols, I., Porter, A.L., Leydesdorff, L.: Science overlay maps: A new tool for research policy and library management. Journal of the American Society for Information Science & Technology 61, 1871–1887 (2010)
6. Leydesdorff, L.: Mapping interdisciplinarity at the interfaces between the Science Citation Index and the Social Science Citation Index. Scientometrics, 391–405 (2007)
7. Porter, A.L., Rafols, I.: Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics 81, 719–745 (2009)
8. Rao, C.: Diversity and dissimilarity coefficients: A unified approach. Theoretical Population Biology, 24–43 (1982)
9. Stirling, A.: A general framework for analysing diversity in science, technology and society. Journal of the Royal Society Interface 4, 707–719 (2007)
10. Rafols, I., Meyer, M.: Diversity and network coherence as indicators of interdisciplinarity: Case Studies in Bionanoscience. Scientometrics 82, 263–287 (2010)
11. Leydesdorff, L., Rafols, I.: A Global Map of Science Based on the ISI Subject Categories. Journal of the American Society for Information Science and Technology 60, 348–362 (2009)
12. Leydesdorff, L., Carley, S., Rafols, I.: Global Maps of Science based on the new Web-of-Science Categories. Scientometrics 81, 1–5 (2012)
13. Batagelj, V., Mrvar, A.: Pajek – Program for Large Network Analysis. Connections 21, 1–11 (1999)
14. Sutherland, R., Eagle, S., Joubert, M.: A Vision and Strategy for Technology Enhanced Learning: Report from the STELLAR Network of Excellence, Bristol, UK, p. 60 (2012)
15. Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A.: How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. Research Policy 41, 1262–1282 (2012)

A Six-Step Guide to Persuasive Learning

Karina Dyrby Kristensen

Department of Communication and Psychology, Aalborg University
Kdkr12@student.aau.dk

Abstract. By combining existing methodological approaches from Persuasive Design and Learning, this paper presents the initial version of a general guide for creating persuasive learning designs. The notion of persuasive learning designs is based on the acknowledgement that there must be an appropriate balance between context and learning technology, and that reflections regarding the context must be included throughout the design process. The 6 step guide to persuasive learning springs from research conducted in the EuroPLOT project, however the approach aims to be generally applicable.

Keywords: Learning Design, Persuasive Design, EuroPLOT project, Persuasive Learning, Context.

1 Introduction

The Six-Step Guide to Persuasive Learning springs from ongoing research in the EU funded research project EuroPLOT, which seeks to explore the theoretical cross field between learning and persuasive design (PD), and to develop learning technologies which is engaging and motivating for the learners to apply. Research within the project has established that the persuasiveness of a technology does not depend solely on its design, but more on the way in which the technology is applied within a given context [1]. The necessity of creating an appropriate balance between the context, the technology and the persuasive intention, in order for the learning design to be persuasive, has proven to be a challenge. In order to address this challenge, this paper presents the initial steps towards a guide for teachers who wish to create persuasive learning designs.

The guide aims to provide teachers with a meta-framework for the development of learning designs, which directs the attention of the teacher towards some of the context oriented reflections which are a requisite for a PD. The guide combines B.J. Fogg's "Eight step design process" for developing persuasive designs [2], with Danish professor in learning and pedagogy Carl Aage Larsen's method for learning designs [3]. Fogg's design model itself takes a wide approach to PD, some of the main messages of the model are to be specific and simple in deciding what the attempted aim of the persuasion is and take the audience abilities into consideration when designing persuasion. Despite Fogg's wide approach, the model lacks the concept of learning. Larsen's approach on the other hand has been criticized from a learning perspective, as focus is placed mainly on the intended outcome and the content, and

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 478–483, 2013.

not enough on the students' abilities [4]. However, when related to PD, Larsen's approach is beneficial as PD by definition focuses on intentional change of attitudes and behaviors [5]. The approach also presents the evaluation part of designing learning as an ongoing process, meaning the teacher's process and decisions will be evaluated as well as the outcome of the learning.

Where this guide clearly differentiates from the two inspirational models is the level of adopting context. As mentioned, it has been established both in EuroPLOT, and in other investigations of PD in relation to other fields [6] that the persuasiveness of a technology is dependent on the design's ability to adapt the intended use context [7]. As such this guide not only combines two existing methods into one approach to persuasive learning, it does so in the acknowledgement that more nuanced understanding of PD is necessary. The guide should be seen as a general tool for creating persuasive learning scenarios.

2 Methods and Materials

This section provides a step-by-step introduction to the guide, with special attention drawn to the elements which are particularly important when creating learning designs which are not only engaging and interactive, but potentially also persuasive.

2.1 Context

When designing persuasive learning it is essential to approach the design process with a constant focus on the intended learning context and to recurrently evaluate and adjust the learning design so that it constantly remains appropriate in relation to the time, place and manner in which the learning activity is to take place. The understanding of context is in PD closely linked to the rhetorical notion of Kairos, which constitutes the opportune moment where a persuasive initiative can be successful [1]. Within this guide, context is perceived in a wide sense, and does as such include an understanding of the physical setting in which the learning design is to take place, as well as a more general understanding of the learning situation and the subject and material which the students are to be presented with. A recurring evaluation and adjustment of the learning design in relation to the context, is essential to the efficiency of the design. All changes made, and all details added both in the overall design of the learning scenario and in the more detailed preparation of learning material, reciprocally influence each other, and as a result may have an impact on the learning designs ability to be persuasive.

Most importantly however, the recurring evaluation should include reflections regarding how the physical settings in the intended context can actively facilitate the intended learning outcome

On the following page, Figure 1 visualizes how The Six-Step Guide to Persuasive Learning distinguishes between the overall design of the learning scenario and the more detailed preparation of learning material. The three first steps which are related to the preparations of learning scenarios reciprocally influence each other, and the order of the steps is not essential. Contrary, the order of the steps is important when proceeding with preparation of learning materials as the specific examples, texts and visualization must facilitate the mediation of primary arguments. All steps are submitted for recurring evaluation and considered in relation to the intended learning context. Once the learning design has been executed it is recommended that a more general evaluation takes place in order for the learning design to be expanded and potentially applied into other domains.

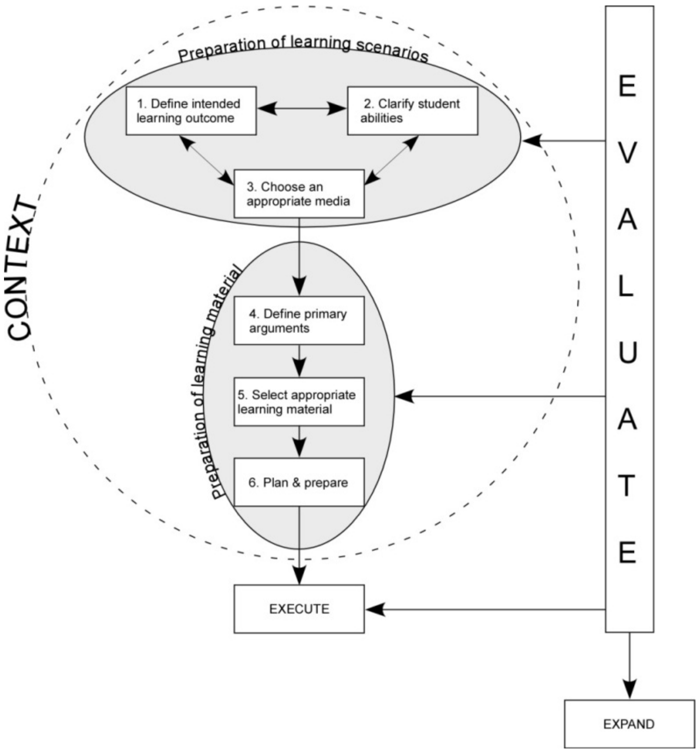


Fig. 1. The Six-Step Guide to Persuasive Learning

2.2 Preparation of Learning Scenarios

Most often, teachers will have an overall idea of the learning scenario they wish to create for their students. Whether the students are at university level or in primary school, teachers will have an intended learning scenario in mind while preparing the learning material. Before proceeding with the individual steps of the guide, it is recommended that the preceding intention with the learning scenario is evaluated in relation to the context, so that the point of origin does not immediately impose a conflict. The learning scenario should facilitate the students' abilities to meet the intended learning outcomes; however it is important that an appropriate balance is maintained between the scenario and the intended learning outcome. Clarifying the overall intention of the learning scenario also constitutes the first primary link to PD, as the intentionality of the design also comprises a core element in this field [1].

Step 1 – Define intended learning outcomes

Unlike the definition of intended learning scenarios, the intended learning outcomes must be specific and simple. Rather than defining an intended learning outcome as “give the student a thorough understanding of poetry theory”, it should be “the student must be able to construct a poem by learned theory”.

A general challenge when defining the intended learning outcomes tends to be that the objectives are too broadly defined. This is an argument which is well known in learning theory [8], but which becomes even more important when the aim is to create persuasive learning designs. Persuasive Technology has do to with changing behavior or attitude, and as such there must be a change from what the students are doing and knowing to where the teacher want them to be afterwards, to conclude if the persuasion was a success. A good rule of thumb in this matter is to ask yourself if what you are trying to achieve is measureable. If not, try to make a new definition which is measureable.

Step 2 – Clarify student abilities

In order to define appropriate learning outcomes, the students’ abilities must be clarified and taken into consideration. Failing to appropriately level the intended learning outcomes to the students preceding abilities will challenge not only the learning outcome but also the persuasiveness of the learning design, as frustrations may distract the students’ attention from the intended outcome.

Furthermore, this step also calls for reflections regarding the media by which the intention is to be communicated (step 3). As the abilities differences in a group the learning outcome should be at a level which challenges the students with less ability and increase the interest for the subject for students with more ability. In some cases it might be an advantage for the teacher to choose a set of different learning outcomes in relation to the different student abilities (step 1).

Step 3 – Choose an appropriate media

The last step of preparation of learning scenarios focuses on the choice of an appropriate media for the persuasive learning design. Based on Fogg’s original perspectives on Persuasive Technology, the media will most often be one or more specific computer technologies. Fogg argues that technologies hold some particular persuasive advantages compared to human persuaders, e. g. they do not get tired of repeating information and they can be time-saving [5]. However persuasive learning does not have to be based solely on a technology. Using a traditional blackboard appropriately may just as well be persuasive depending on the context. It may however be most beneficial to consider a combination of media which facilitates each other and enables the students to actively engage with the learning material rather than remain passive and observant. Most importantly when creating persuasive learning designs however is to choose an appropriate media where both the student and the teacher feel familiar and comfortable. By enabling students to learn within their comfort zone the media will not draw attention away from the intended outcome, but rather make it more simple and easy to grasp.

2.3 Preparation of Learning Material

Once the general understanding of the intended learning outcome has been defined, this second layer of the guide draws the attention towards more specific details regarding the preparation of learning material. This part of the design process will naturally incorporate more elements from traditional approaches to learning design and the steps described in this paper should as a result be expanded by the teacher as the design process moves forward.

Step 4 – Define primary arguments

Persuasion is often linked to classical rhetoric [9], and as a result, the development of persuasive learning designs has much to learn from the art of logical argumentation. The primary arguments serve as a guide for the teacher both during the execution of the learning, but also in the preparation of the learning, and should be seen as a core element in the learning design. The primary arguments help to define the appropriate learning material (step 5), they support and underline the intended learning outcome (step 1), and they constitute the material which should be presented via the chosen media (step 3).

Step 5 – Select appropriate learning material

Once the primary arguments are defined, learning material such as texts, illustrations, film clips etc. can be selected in order to best provide a rich presentation of each individual argument. - Still keeping in mind that the arguments and the specific learning material must be appropriately balanced against both the context and the elements in step 1-3. The learning material will most likely require adjustments if the learning design is to be applied to different groups of students.

Step 6 – Plan & prepare

In the final step, the preparation of the course and its execution is finalized. The decisions made in the earlier steps are summed up, evaluated and adjusted. Also, it should be considered if other persuasive strategies such as rewards should be included in the learning design.

3 The Three E's; Execute, Evaluate and Expand

Although reflections regarding the context play a vital role in the development of persuasive learning designs, the most important activities are in fact those which take place once the process moves beyond the contextual reflections. Once step six is completed, the persuasive learning design is ready for execution. Another primary overlap between PD and learning comprises the fact that it is once the student (or user) interacts with the design that persuasion may potentially take place. It is also in the execution of the learning design that the teacher is able to evaluate its effect.

Execution of a persuasive learning design should always be followed up by an overall evaluation of not just the design but also the learning situation as a whole. Evaluation should involve not only the designing teacher, but also the students and any potential observers. Evaluating a design in order to make improvements is always beneficial, but when implementing elements from PD, the evaluation becomes even more vital as the line between persuasion and manipulation is delicate, and ethics is argued to constitute a core element in the unique claim of the field [7].

Learning in general calls for a number of objectives and milestones. By following this guide only one or potentially a few of these objectives may be reached, due to the previously described simplicity of the intended learning outcome. When all aspects of that process has been evaluated, difficulties been corrected and the end result, the intended learning outcome defined in step one, is reached. The time has come for a possible expansion on the success, either in terms of adjusting the complexity of the existing design, or by moving on to developing new persuasive learning designs.

References

1. Gram-Hansen, S.B., Persuasive Learning Design Framework. EuroPLOT Deliverable, vol. 3.3. Aalborg Universitet, Aalborg (2012)
2. Fogg, B.J.: Creating Persuasive Technologies: An Eight-Step Design Process. In: Proceedings of the 4th International Conference on Persuasive Technology. ACM, Claremont (2009)
3. Larsen, C.A., Høeg Larsen, C.A.: Didaktiske emner belyst gennem 12 artikler. Danmarks Pædagogiske Bibliotek, Århus (2002)
4. Hiim, H., Hippe, E.: Læring gennem Oplevelse, Forståelse og Handling. Gyldendalske Boghandel, København (2007)
5. Fogg, B.J.: Persuasive Technology - Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco (2003)
6. Lykke, M.: Persuasive design strategies: Means to improve the use of information organisation and search features in web site information architecture? In: Proceedings ASIST Special Interest Group on Classification Research 20th Workshop. ASIST, Vancouver (2009)
7. Gram-Hansen, S.B., Gram-Hansen, L.B.: On the role of Ethics in Persuasive Design. In: ETHICOMP 2013. SDU, Kolding (2013)
8. Biggs, J., Tang, C.: Teaching for Quality Learning at University, 4th edn. McGraw-Hill and Open University Press, Maidenhead (2011)
9. Christensen, A.-K.K., Hasle, P.F.V.: Classical Rhetoric and a Limit to Persuasion. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 307–310. Springer, Heidelberg (2007)

Scaling Informal Learning: An Integrative Systems View on Scaffolding at the Workplace

Tobias Ley[1], John Cook[2], Sebastian Dennerlein[3], Milos Kravcik[4], Christine Kunzmann[5], Mart Laanpere[1], Kai Pata[1], Jukka Purma[6], John Sandars[7], Patricia Santos[2], and Andreas P. Schmidt[8]

[1] Tallinn University, Estonia
{tley,kpata,martl}@tlu.ee
[2] University of the West of England
{john2.cook,Patricia.Santosrodriguez}@uwe.ac.uk
[3] Graz University of Technology, Austria
sdennerlein@tugraz.at
[4] RWTH Aachen University
kravcik@dbis.rwth-aachen.de
[5] Pontydysgu, UK
kontakt@christine-kunzmann.de
[6] Aalto University, Finland
jukka.purma@aalto.fi
[7] University of Leeds, UK
J.E.Sandars@leeds.ac.uk
[8] Karlsruhe University of Applied Sciences, Germany
andreas_peter.schmidt@hs-karlsruhe.de

Abstract. While several technological advances have been suggested to scale learning at the workplace, none has been successful to scale informal learning. We review three theoretical discourses and suggest an integrated systems model of scaffolding informal workplace learning that has been created to tackle this challenge. We derive research questions that emerge from this model and illustrate these with an in-depth analysis of two workplace learning domains.

Keywords: Workplace Learning, Informal Learning, Scaffolding, Scaling.

1 Introduction

Despite the recognized importance of informal learning at the workplace, most technological solutions are targeted towards a learning model based on the ideas of direct formal instruction [1]. In contrast to formal instruction that is organised along curricula, workplace learning takes place through work processes, is multi episodic, is often informal, is problem based and takes place on a just in time basis [2] and often involves the passing on of skills and knowledge from skilled workers [3]. Learning trajectories [4] emerge as a result of those learning episodes in the context of situated learning, when individuals meaningfully connect them, reflect about them and advance their competence.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 484–489, 2013.

While this learning from individual experience is highly effective and intrinsically motivating for the individual [6], it does not scale very well: if individual experiences are not further taken up in systematic organizational learning practices, learning remains costly, fragmented and unsystematic. Scaling up informal learning would enable a learner to receive meaningful, relevant and individualized support for his learning needs in the context of his work, and take better advantage of the multitude of learning opportunities that are available around him. A number of learning technologies have been suggested to scale learning, but each focus only on a single aspect of informal learning. *Adaptive Learning Technologies* scale guidance by codifying some of the strategies and rules that a human tutor would use. *Social Networking Technologies* scale personal interactions by extending and augmenting the reach of personal network. *Semantic Technologies* scale the representation and generation of meaning.

With this paper, we suggest an integrative model for scaffolding informal workplace learning that integrates these technological perspectives, and suggests new research directions to scale informal learning at the workplace.

2 Towards an Integrated View of Scaffolding Informal Learning

The technologies briefly discussed in the previous section have each originated from one of three distinct theoretical discourses on learning and its support (see Figure 1). The first of these is the *Pedagogical Perspective* which deals with scaffolding self-regulated learning. Scaffolding as a metaphor refers to the provision of temporary support for the completion of a task that a learner might otherwise be unable to achieve [7]. This perspective, therefore, looks at the effects of temporary support structures on learning with the aim of facilitating self-regulative processes. Scaffolding is an approach to providing relevant guidance for learning by grounding the task between a more capable peer or teacher and the learner, thus creating a shared understanding of the task [7,8]. This requires fine-tuned support based on an ongoing diagnosis of the learner's level of understanding and changing knowledge and skills [9]. There is a close relationship to adaptive learning technologies.

Whereas the pedagogical perspective puts a focus on the individual agent's learning, scaffolding in socio-technical systems is not restricted to interactions between individuals with differing skill levels, but it also includes interactions with artefacts, networks and peer groups [10]. A Community of Practice (CoP) [11] is a concept to systemize these interactions. CoP develop as a learning collectivity in situated workplace setting, in which persons have dense relations of mutuality based on social relationships (for example trusting each other based on expertise and support) and social bonds (based on working with each other or having virtual connections). Nardi et al [12] highlight the use of embedded intensional networks which are highly strategic personal networks to meet individual learning needs. Within the CoP, members share cognitive communality which is created while working on joint enterprise, using shared repertoire (tools, objects, artifacts, rules) and shared knowledge. Learning in a CoP happens by a process of peripheral legitimate participation in which newcomers are encultured to the CoP. In extending CoP, the knowledge maturing framework [13] takes a closer look at how collective knowledge is developed along a number of discrete phases, and how the characteristics of knowledge change. We call this perspective the *Community Perspective*.

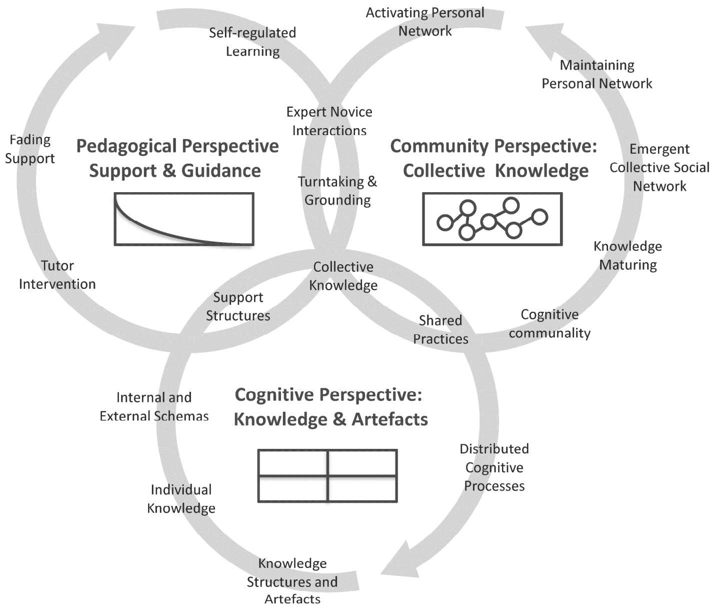


Fig. 1. An integrated systems view on scaffolding informal learning at the workplace

Finally, a *Cognitive Perspective*, more clearly specifies how human performance in the workplace and its support relies on cognitive schemas. Schemas are knowledge structures, in which typical relations of the reality are represented based on previously made experiences reflecting a common way to make sense out of experiences, to represent typical tasks or problems [14]. Cognitive schemas possess a procedural component, in which bottom up and top down processes interact: i.e. specific environmental information activate a certain schema, which then determines further actions. Schemas may have conscious element during learning, but over time, schema application and use is automatized and enhances expert performance. Besides their function as internal knowledge representations, parts of schemas can also be explicated and represented externally in artifacts. In distributed networks of actors and artifacts that interact in a meaningful way, some knowledge is represented only in a distributed cognitive system [15]. Distributed cognition occurs when one cognitive task (e.g. solving a problem or making a decision) is distributed among people and artifacts that act as a single system [18] (e.g. in workflows and production chains).

Each perspective (Pedagogical, Community and Cognitive) looks at a system that in itself is self-contained (each of the three circles). For example, in the pedagogical perspective, a self-regulated learner enters into interaction with a more capable peer. Through grounding, the two negotiate a common understanding of the task, and the more capable peer then uses support structures to adapt the support to the learner, transferring more and more responsibility to the learner. This in turn should influence self-regulative competence as well as acquisition of domain knowledge.

While each of the three systems is operationally closed and can therefore be analysed independently, we suggest that interactions between these sub-systems happen through structural coupling (see [19] for a similar argument) and this connects different levels of analysis: processes taking place in one of the systems trigger effects in another system. It is in these intersections where new research questions emerge.

3 Research Areas and Examples from Two Domains

From the discussion in the previous section, the following research areas emerge:

- Community and Pedagogical: How does collective knowledge emerge and mature in a community setting, and is then utilized in individual scaffolding interactions?
- Community and Cognitive: How do individual and collective knowledge influence each other, and how are these represented in digital or physical artefacts?
- Cognitive and Pedagogical: How do people appropriate and make sense of distributed knowledge representations and how are these utilized in scaffolding?

We have recently set out to tackle those questions[1]. Next, we illustrate our approach with results from an in-depth analysis of two workplace learning domains.

3.1 Health Care: Scaffolds Emerge from Collective Knowledge Processes

In the UK, health sector national guidelines are published by NICE (http://www.nice.org.uk/) in three areas: the use of health technologies, clinical practice and guidance for public sector workers on Health promotion and ill-health avoidance. The guidelines are a result of considering all available scientific evidence to be taken into account when making decisions about treating a patient (evidence-based medicine). They are interpreted locally by General Practitioners (GPs) and other Health workers, and used in local practices. This local interpretation is challenging because the guidelines are written for the average and most common cases. In actual practice, however, GPs deal with individual patients with very specific conditions where it is sometimes necessary to deviate from the guidelines. In such cases, they may seek validation from a colleague in their trusted network. This can be considered a much weaker form of guidance and scaffolding as it revolves around the discussion of individual cases. Social network technologies could make these trusted personal networks more readily available for the person seeking validation and lead to "living local guidelines" that are interpreted in practice and adapted to local conditions. How trust emerges and is represented in such network settings is one of the questions that we set out to answer.

---

[1] Project "Learning Layers - Scaling up Technologies for Informal Learning in SME Clusters" (`http://www.learning-layers.eu`), partially funded under FP7, grant no: 318209

In the wider professional network these discussions that evolve around the existing guidelines could be accumulated using semantic technologies to show where adaptations to the guidelines could be necessary. For example, certain rare illnesses are being introduced by increased foreign travelling and become more common. As these are being discussed, this emergent collective knowledge can be used to scaffold learners before these conditions are actually introduced into the guidelines.

3.2 Building and Construction: Scaffolding in Distributed Cognitive Systems

Advanced educational institutions in the German construction sector require apprentices to complete several projects alongside their further training. These projects always follow the same sequence of steps learners have to go through: preparation, execution and quality assessment. This structure has developed over years as part of the practice of the construction trade and has been transferred into formal learning processes. While experts have internalized these structures and follow them automatically, apprentices learn them by keeping a paper-based portfolio (called "white folder") of their training projects, where each project is structured in the same way.

Digitizing the "white folder" would enable apprentices to easily return to some of their learning when back at the construction site of their employers. There they soon discover that at the real construction site, things often deviate from the taught standards. This provides the starting point for informal learning situations and enrichment of formal learning experiences. The digital white folder could enable the learners to co-edit and share knowledge on projects and to exchange their experiences. Thereby a distributed cognitive system could be created through which learning can be scaffolded. How individual knowledge and collective knowledge interact in such situations, i.e. how learners appropriate the aggregated experiences of others into meaningful assemblies, and how this can be adequately supported will be a focus of the project.

A particular challenge in building and construction is the role that physical tools and materials play in the learning process. While there is a body of knowledge about the use of tools and materials that can be taught in educational institution, there is also a large amount of individual knowledge among workers about specific tools and properties of construction sites. If some of this knowledge can be captured into related objects and locations, these objects and locations can work as improved support structures for apprentices. This could happen through tagging physical objects with temporary, company- or site-specific information about their actual use. We are testing how various wearable recording devices are suitable to record the use of objects, and are developing mobile tools to annotate and edit the recordings into learning resources.

4 Conclusions and Outlook

Current technologies mostly facilitate formal learning in well-structured domains. To scale informal learning in complex and dynamic domains, we have suggested taking a systems perspective that integrates perspectives on individual actors, their cognition and interactions within a system of actor networks and communities. This perspective eventually suggests viewing scaffolding as an adaptive system, in which the individual learner and the network adapt to each other in the scaffolding process.

References

1. Kraiger, K.: Transforming our models of learning and development: Web-based instruction as enabler of third-generation instruction. Industrial and Organizational Psychology 1(4), 454–467 (2008)
2. Kooken, J., Ley, T., de Hoog, R.: How Do People Learn at the Workplace? Investigating Four Workplace Learning Assumptions. In: Duval, E., Klamma, R., Wolpers, M. (eds.) EC-TEL 2007. LNCS, vol. 4753, pp. 158–171. Springer, Heidelberg (2007)
3. Attwell, G., Baumgartl, B., Balica, M., Figueira, E., Grollmann, P., Kämäräinen, P., van Elk, L.: Creating Learning Spaces: Training and Professional Development for Trainers. Navreme, Vienna (2008)
4. Eraut, M.: Editorial. Learning in Health and Social Care 4(1), 1–6 (2005)
5. Attwell, G.: New roles for vocational education and training teachers and trainers in Europe: A new framework for their education. J. of Europ. Industrial Training 21(6/7), 256–265 (1997)
6. Hague, C., Logan, A.: A review of the current landscape of adult informal learning using digital Technologies. General Educators Report (2009)
7. Wood, D., Bruner, J.S., Ross, G.: The role of tutoring in problem solving*. J. of Child Psychology and Psychiatry 17(2), 89–100 (1976)
8. Collins, A., Brown, J.S., Newman, S.E.: Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In: Resnick, L.B. (ed.) Knowing, Learning, and Instructions: Essays in Honor of Robert Glaser, pp. 453–494. Erlbaum (1989)
9. Azevedo, R., Hadwin, A.F.: Scaffolding self-regulated learning and metacognition–Implications for the design of computer-based scaffolds. Instructional Science 33(5), 367–379 (2005)
10. Puntambekar, S., Hubscher, R.: Tools for Scaffolding Students in a Complex Learning Environment: What Have We Gained and What Have We Missed? Educational Psychologist 40(1), 1–12 (2005)
11. Wenger, E.: Communities of practice: Learning as a social system. Systems Thinker 9(5), 2–3 (1998)
12. Nardi, B.A., Whittaker, S., Schwarz, H.: NetWORKers and their activity in intensional networks. Computer Supported Cooperative Work 11(1-2), 205–242 (2002)
13. Schmidt, A., Hinkelmann, K., Ley, T., Lindstaedt, S., Maier, R., Riss, U.: Conceptual Foundations for a Service-oriented Knowledge and Learning Architecture: Supporting Content, Process and Ontology Maturing. In: Pellegrini, T., Auer, S., Tochtermann, K., Schaffert, S. (eds.) Networked Knowledge - Networked Media. SCI, vol. 221, pp. 79–94. Springer, Heidelberg (2009)
14. Mandl, H., Spada, H.: Wissenspsychologie. Psychologie Verlags Union (1988)
15. Halverson, C.A.: Activity theory and distributed cognition: Or what does CSCW need to DO with theories? Comp. Supported Coop Work 11(1-2), 243–267 (2002)
16. Vosniadou, S.: The cognitive-situative divide and the problem of conceptual change. Educational Psychologist 42(1), 55–66 (2007)
17. Rasmussen, J.: The importance of communication in teaching: A systems-theory approach to the scaffolding metaphor. J. of Curriculum Studies 33(5), 569–582 (2001)

Active Learners: Redesigning an Intelligent Tutoring System to Support Self-regulated Learning

Yanjin Long and Vincent Aleven

Human Computer Interaction Institute, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
{ylong,aleven}@cs.cmu.edu

Abstract. Supporting students' self-regulated learning (SRL) is an important topic in the learning sciences. Two critical processes involved in SRL are self-assessment and study choice. Intelligent tutoring systems (ITSs) have been shown to be effective in supporting students' domain-level learning through guided problem-solving practice, but it is an open question how they can support SRL processes effectively, while maintaining or even enhancing their effectiveness at the domain level. We used a combination of user-centered design techniques and experimental classroom research to redesign and evaluate an ITS for linear equation solving so it supports self-assessment and study choice. We added three features to the tutor' Open Learner Model (OLM) that may scaffold students' self-assessment (self-assessment prompts, delaying the update of students' progress bars, and providing progress information on the problem type level). We also designed a problem selection screen with shared student/system control and game-like features. We went through two iterations of design and conducted two controlled experiments with 160 local middle school students to evaluate the effectiveness of the new features. The evaluations reveal that the new OLM with self-assessment support facilitates students' learning processes, and enhances their learning outcomes significantly. However, we did not find significant learning gains due to the problem selection feature. This work informs the design of future ITS that supports SRL.

Keywords: Self-assessment, study choice, intelligent tutoring system, open learner model, user-centered design, classroom evaluations.

1 Introduction

Intelligent tutoring systems (ITSs) typically focus on supporting domain-level learning and have been illustrated to be effective at doing so [5]. Researchers are now more interested in how ITS can be designed to foster self-regulated learning (SRL), while maintaining or even enhancing their effectiveness in supporting domain-level learning. Self-assessment and study choice are two critical processes involved in SRL. Self-assessment refers to monitoring and evaluating how well you are learning/have learned. Study choice means the learner selects what s/he will work on next during the learning process. Feyzi-Behnagh and colleagues [2] found that the metacognitive

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 490–495, 2013.

prompts and feedback provided by the pedagogical agents in MetaTutor could enhance students' self-assessment accuracy and learning efficiency (but not the learning effectiveness) [2]. However, with such promising results from previous work, it is still an open question how we can design specific features of the ITSs to support self-assessment and study choice effectively, while maintaining or even enhancing their effectiveness at the domain level. Recently, ITS researchers have pointed out the potential of using Open Learner Models (OLMs) to support students' self-assessment and study choice [1]. OLM is one component of ITS that displays students' learning progress using different forms of visualizations. In the current work, we explore new designs of an ITS (especially the OLM) to identify the features that can effectively scaffold SRL processes (especially self-assessment and study choice), which will also lead to improved domain-level learning outcomes. We went through design and evaluation stages including paper prototyping, high fidelity prototyping, building redesigned tutor version 1, classroom experiment 1, building redesigned tutor version 2 and classroom experiment 2. Throughout the whole design process, we combined different research approaches, including HCI/user-centered design techniques, experimental educational research and educational data mining. We describe the methods and results used at each stage of the design process to articulate the rationales for our designs, and discuss insights for future work.

2 Design Process, Methods and Results

2.1 Paper Prototyping and High Fidelity Prototyping

We used a tutor for solving linear equations as the platform for this study. The tutor is an example-tracing tutor built with Cognitive Tutor Authoring Tools [5]. It provides practice for five types of linear equations. The example-tracing tutors have built-in OLMs (the Skillometer), which shows students' progress in the form of skill meters.

As the first step in user-centered design, in order to find out the needs of the users, we conducted an interview study with 44 high school students. The interviews revealed that the students inspected the OLM frequently to see their progress, but thought the design of the OLM was too simple to convey much progress information. They did not actively reflect or self-assess in the tutor either. Besides, the students expressed strong interests in selecting their own problems in the tutor. Based on the results from this study, we decided to redesign the Skillometer and create a separate screen for problem selection. We built both paper and high fidelity digital prototypes that show different screens and alternative designs. We conducted two rounds of user testing through one-on-one think aloud sessions with 7 middle school students. The main features of the new designs aimed to 1) facilitate self-assessment; 2) provide more complete/multi-level progress information to the students; and 3) give them control over problem selection. We gathered three primary design recommendations:

1) Facilitate self-assessment using explicit prompts. During the think alouds, the students did not actively initiate any self-assessment activities. To facilitate self-assessment in the tutor, we need to add explicit self-assessment prompts/questions.

2) Gamify by creating levels. We tried to provide more complete progress information by showing students' overall progress on the five different levels (problem types) in addition to the skill bars, which also adds elements of gamification to the system. The students really liked this game-like feature. The participants also expressed that displaying the progress of each problem type on the same screen where they have to select the next problem was helpful for them to make decisions.

3) Share control over problem selection between student and system. All participants admitted that they might keep selecting easy problems if they were completely free to select problems by themselves. Therefore, to prevent such suboptimal problem selection decisions, also to maintain the effectiveness of mastery learning in the system, we decided that once the system deems that the student has reached mastery for a certain level, they can no longer select new problems from that level. Such joint control with the system grants students freedom but prevents them from abusing the system.

2.2 Redesigned Tutor Version 1 and Classroom Experiment 1

We then implemented a fully-functional version of the tutor with support for self-assessment and study choice, based on our prototypes. This redesigned tutor has four key new features:

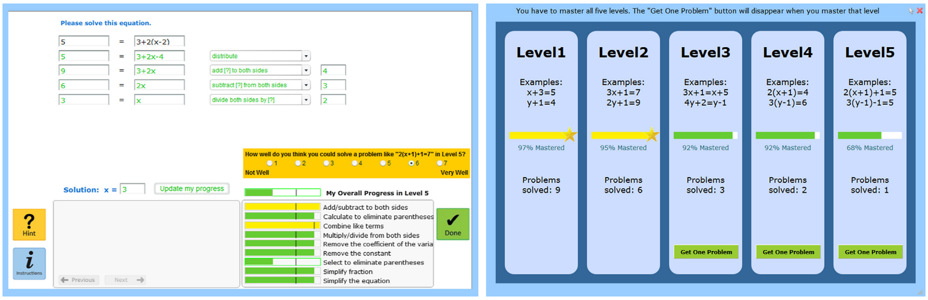


Fig. 1. The redesigned problem solving screen (left) and the problem selection screen (right)

1) Delaying the update of the progress bars. Instead of updating the bars while the student is in the midst of solving a problem, the new tutor updates the bars only when the student is done with the problem, so as to make it easier for students to focus their attention on the changing of the bars. The "Update my progress" button appears after the student finishes solving the problem. When the student clicks on it, the tutor updates the bars (i.e., the bars move to their new positions, based on the students' performance on the problem they just completed). The black lines marked on the bars allow a before/after comparison to further emphasize the change of progress.

2) Self-assessment prompt. The self-assessment prompt ("How well do you think you can solve a problem like "" in level x") appears after the bars stop updating (shown in orange on the left in Figure 1). Answering this self-assessment question may help students become better at self-assessing and self-assess more actively. Also, given the self-assessment prompt comes just prior to the problem selection screen, answering it may help them make better problem selection decisions.

3) Showing progress on the problem type level. Showing both the detailed progress on the skill level and the overall progress on the problem type level gives students more complete information regarding their learning and may further support useful reflection and self-assessment.

4) Selecting the next problem. As shown on the right in Figure 1, on the problem selection screen students can view and compare their progress on different levels, which can aid them in deciding which level they want to work on next. Once they click any "Get One Problem" button, they are directed to a new problem solving screen with a problem from the chosen level. Students can only select problems from unmastered levels, which prevents one type of suboptimal problem selection decision students might make (selecting problems that are too easy), while still giving them some freedom and control over their own learning.

Experimental Design, Participants and Procedure of Classroom Experiment 1. To empirically evaluate the effectiveness of the new tutor, we decided to conduct a controlled experiment instead of a traditional HCI user study. We believe the controlled experiment can provide rigorous evidence as to whether/how our new designs make a difference in student learning. We conducted a 2x2 experiment with independent factors OLM (whether or not the OLM (both skill-level and problem type-level of progress information) are shown to the students) and PS (whether or not students could select their next problem from an unfinished level) at a local public middle school. 98 8^{th} grade students were randomly assigned to one of the four conditions. This is an ablation experiment and we modified the interfaces accordingly to match the manipulation. All participants completed a paper pre-test on solving the five types of linear equations, worked with the tutor for three class periods, and completed a paper post-test that was in the same format as the pre-test.

Results. There were no statistically significant differences among the four conditions either on the pre-test or the post-test. Also, no significant improvements from pre- to post-tests were found for the students. However, analysis of the process measures from tutor log data reveals some promising benefits of having access to an OLM during learning in the tutor. Students who learned with an OLM needed fewer problems to reach mastery on each level ($F(1, 435) = 4.450$, $p = .035$, $\eta^2 = .010$), made fewer incorrect attempts when solving each problem ($F(1, 435) = 4.922$, $p = .027$, $\eta^2 = .011$), and needed less help in the tutor (based on the average assistance score ((hints + incorrect attempts) / total steps) per problem, $F(1, 435) = 6.557$, $p = .011$, $\eta^2 = .015$). On the other hand, the students who had the freedom of problem selection asked for more hints than students who did not have problem selection ($F(1, 435) = 5.642$, $p = .018$, $\eta^2 = .013$).

Discussion. We had a ceiling effect on pre-test with the sample in this experiment, and did not find any significant learning gains due to the tutor. However, the analysis of the tutor log data shows some benefits of having access to an OLM with support for self-assessment. The students who had the OLM needed fewer problems to reach mastery, made fewer incorrect attempts, and needed less assistance in the tutor. These results suggest that the OLM with self-assessment support facilitated reflection and self-assessment during learning, which led to more efficient learning process. Nevertheless, given the small effect sizes (η^2 is around .01) of these log data analysis, it is reasonable to strengthen the scaffolding for self-assessment to foster stronger effects. Therefore, we further improved our design of the OLM and conducted a new controlled experiment with younger students (7^{th} grade).

2.3 Redesigned Tutor Version 2 and Classroom Experiment 2

We kept the four new features in version 1, but further revised the design of the OLM on the problem solving screen, aiming to strengthen the scaffolding for self-assessment and reflection, so that students can achieve better learning outcomes.

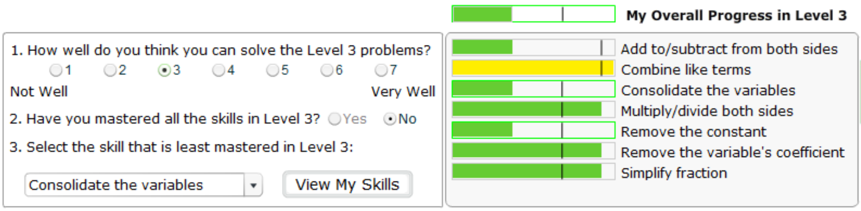


Fig. 2. The new OLM view on the problem solving screen

1) Add more specific self-assessment prompts. We further guide students' self-assessment by adding two more specific self-assessment prompts that are tied directly to their skills. The two new prompts are "Have you mastered all the skills in Level x" and then asking the students to select the least mastered skill in that level from a drop-down menu (as illustrated on the left side in Figure 2).

2) Hide the progress information until the self-assessment questions are answered. In the redesigned tutor version 1, the students answered the self-assessment prompt after the progress bars had been updated to reflect performance on the last problem. It is possible that students answered the prompts based just on what their bars look like, without much reflection. Possibly, students would reflect more strongly on their skills if they self-assess before the skill bars are updated and shown to them. Therefore, we updated the problem solving screen so the bars are hidden until after the student has worked through the self-assessment prompts (the three questions shown on the left in Figure 2; the progress bars on the right in Figure 2 are initially hidden). After students answered all three self-assessment prompts, the "View My Skills" button appears. Once students click the "View My Skills" button, the level and skill bars are shown and start updating after 1 second. In this way, the updating of the skill bars serves as a form of instant feedback on students' self-assessment.

The problem selection screen remained unchanged. With the revised tutor version, we conducted a new controlled experiment with 62 7th grade students at another local public school. The experimental design, procedure and measurements were the same as experiment 1. The results of this experiment have been reported in another paper [4]. To summarize, we found that all students' knowledge of solving linear equations improved significantly from pre- to post-tests, affirming the effectiveness of the tutor. More interestingly, we found that having access to an OLM resulted in better performance on the post-test, which was also supported by log data analysis. However, we did not find significant effects of the PS factor. The results from the new study affirm the effectiveness of the OLM on students' domain level learning, while the effects of problem selection still need further investigation.

3 Conclusions and Future Work

This paper documents and describes the iterative process we went through to redesign a linear equation tutor so it supports self-assessment and study choice, key processes in self-regulated learning. By scaffolding these processes, we seek to improve students' domain-level learning (and did so for supporting self-assessment). Through this design process, we identified three key features of the OLM that could scaffold students' self-assessment and reflection: self-assessment prompts, delaying the update of the progress bars and providing progress information on the problem type level. In reflection, we believe that the combination of HCI techniques and quantitative educational research methods is an effective way of exploring open-ended design questions in educational technologies. The two approaches weave together and work well in generating design ideas, iteratively improving the designs, and rigorously evaluating the design products. We plan to extend this work by exploring designs to support other SRL processes in the tutor, such as goal setting. Eventually, we hope to help students become better self-regulated learners, who are active and efficient in planning, monitoring and evaluating their learning.

Acknowledgements. We would like to thank the participating teachers and students. This work is funded by an NSF grant to the Pittsburgh Science of Learning Center (NSF Award SBE0354420).

References

1. Bull, S., Ahmad, N., Johnson, M., Johan, R., Mabbott, A., Kerly, A.: Adaptive Navigation Support, Learner Control and Open Learner Models. In: Nejdl, W., Kay, J., Pu, P., Herder, E. (eds.) AH 2008. LNCS, vol. 5149, pp. 275–278. Springer, Heidelberg (2008)
2. Feyzi-behnagh, R., Khezri, Z., Azevedo, R.: An Investigation of Accuracy of Metacognitive Judgments during Learning with an Intelligent Multi-Agent Hypermedia Environment. The Annual Meeting of the Cognitive Science Society, 96–101 (2011)
3. Long, Y., Aleven, V.: Students' Understanding of Their Student Model. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) AIED 2011. LNCS, vol. 6738, pp. 179–186. Springer, Heidelberg (2011)
4. Long, Y., Aleven, V.: Supporting Students' Self-Regulated Learning with an Open Learner Model in a Linear Equation Tutor. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS, vol. 7926, pp. 219–228. Springer, Heidelberg (2013)
5. Waalkens, M., Aleven, V., Taatgen, N.: Does Supporting Multiple Student Strategies Lead to Greater Learning and Motivation? Investigating a Source of Complexity in the Architecture of Intelligent Tutoring Systems. Computers & Education 60, 159–171 (2013)

Answering Confucius: The Reason Why We Complicate

Bernardo Pereira Nunes[1,2,3], Stella Pedrosa[3], Ricardo Kawase[2], Mohammad Alrifai[2], Ivana Marenzi[2], Stefan Dietze[2], and Marco Antonio Casanova[1]

[1] Department of Informatics - PUC-Rio - Rio de Janeiro, RJ - Brazil
{bnunes,casanova}@inf.puc-rio.br
[2] L3S Research Center, Leibniz University Hannover, Germany
{nunes,kawase,alfrai,marenzi,dietze}@l3s.de
[3] Central Coordination for Distance Learning - PUC-Rio - Rio de Janeiro, RJ - Brazil
{bernardo,stella}@ccead.puc-rio.br

Abstract. Learning is a level-progressing process. In any field of study, one must master basic concepts to understand more complex ones. Thus, it is important that during the learning process learners are presented and challenged with knowledge which they are able to comprehend (not a level below, not a level too high). In this work we focus on language learners. By gradually improving (complicating) texts, readers are challenged to learn new vocabulary. To achieve such goals, in this paper we propose and evaluate the 'complicator' that translates given sentences to a chosen level of higher degree of difficulty. The 'complicator' is based on natural language processing and information retrieval approaches that perform lexical replacements. 30 native English speakers participated in a user study evaluating our methods on an expert-tailored dataset of children books. Results show that our tool can be of great utility for language learners who are willing to improve their vocabulary.

Keywords: Technology enhanced learning, language development, learning process.

1 Introduction

Reading is a fundamental activity for all areas of knowledge. The practice of reading, strongly linked to the learning process, starts in the early school years and remains throughout life. Although, it requires mastery of certain techniques, reading is not a technical competence, but a process that begins in the relationship between the reader and text and continues by making sense of the text and promoting the development of new ideas influenced by prior background. For instance, reading helps to develop vocabulary in various forms of written and oral expressions. Thus, reading precedes writing being the main provider of basic elements for the production of texts.

A text ends up from the reading and the meaning that the reader gives to the text according to his understanding and the associations made shaped by the reader's prior knowledge and experiences. In this manner, it is possible to consider reading as a dialogic attitude in which the reader triggers threads of thoughts from a set of relationships with the text.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 496–501, 2013.

Trying to reduce the cognitive overload of reading activities, we often seek to develop activities that involve the simplification of texts, i.e., simplification of text preserving its original meaning [4]. Conversely, little has been explored about the possibilities of the *text sophistication* can make towards developing vocabulary. The introduction of words that are unusual for an individual or a group is an opportunity to expand their vocabulary knowledge. This can be achieved by transforming a text with simple vocabulary, read and discussed beforehand, into a more sophisticated terminology focusing on an individual or group.

The language develops in experiences influenced by sociocultural surroundings. If, due to various factors, this environment offers limited opportunities, the vocabulary of this group will be restricted and phrasal structures will be simple. Previous studies show that the language development of children is related to the sociocultural environment and that school interventions in early childhood education can minimize the differences between these children and those included in a privileged sociocultural environment with a greater range of opportunities. Thus, it imposes a natural limit to the expansion of vocabulary. However, reading nurtures new experiences and opportunities that contribute in the process of vocabulary acquisition and language development. This trend, coordinated with other activities, allows the *acquisition*, *expansion*, and *formation* of a more complex vocabulary, which contributes to learning any language, native or not.

In particular, communication is undeniably relevant in social relations amongst the groups that individuals attend. The vocabulary of a group gives the individual a sense of belonging and the sociocultural migration that education can provide is often barred, or at least hampered, by the limitations of the acquisition of new vocabulary. We do not suggest that expressions that are part of the sociocultural environment of origin should be overlooked, but they could be added to allow the expansion of types of communication and other sociocultural contexts. Another reason for vocabulary development is to improve communication. In the classroom, teaching vocabulary is often overlooked, although it is of well-known importance for learners of foreign languages to express their ideas clearly. Thus, new learning strategies are needed for learning vocabulary and development of autonomy.

In this paper, we introduce the 'complicator', a method that construes given sentences into a more sophisticated vocabulary. As hereby mentioned, the rationale behind the method is that one can learn (improve his vocabulary) by reading sentences that contain new and infrequent terms.

2 Method

In this section, we present our method for text sophistication based on lexical replacements as depicted in Figure 1. The method is divided into 4 main steps: (i) part-of-speech (POS) tagging; (ii) synonym probing; (iii) context frequency-based lexical replacement; and (iv) sentence checker.

2.1 Part-of-Speech Tagging

Words cannot be exchanged disregarding the context, otherwise, the sentence may result with a different meaning. For this reason, a first step is to identify the right part-of-speech of each word in a sentence and then look for a suitable synonym.

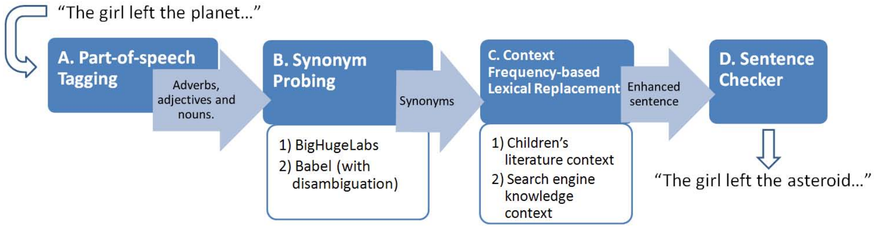


Fig. 1. Complicator workflow

In this step we used the Stanford Log-linear Part-Of-Speech (POS) Tagger [5] to annotate each word in a given sentence. We focused on 3 POS groups: *adjectives*, *nouns* and *adverbs*. As depicted in Figure 1, the first step outputs the original sentence with the POS tags assigned to each word.

2.2 Synonym Probing

In this step, we identify synonyms of given nouns, adverbs and adjectives of a sentence. As a word can have multiple synonyms, we divide the synonyms into categories and filter out the synonyms that express different meanings.

To identify the sense of a word in a given context, we used the Babelnet API[1] developed by Navigli and Ponzetto [3]. Babelnet is built on the top of Wordnet[2], which is a lexical database in English that groups words into synsets, i.e. words that denote the same concept.

We also used a thesaurus database[3] of the Big Huge labs. However, this thesaurus does not provide any information regarding the sense of each word. Hence, we only filter out the synonyms that belong to a different POS category and without filtering by sense.

Thus, this step is responsible for finding a set of synonyms for a particular word in the original sentence and for outputting a filtered set of synonyms according to a specific context.

2.3 Context Frequency-Based Lexical Replacement

As we already have the set of synonyms filtered by sense from the previous steps, this step aims at identifying a synonym for a word that best fits in a determined context. For this, we rely on how frequent a word is found in a controlled vocabulary. The synonym that will replace the original word is the next most frequent word in the vocabulary, but less frequent than the original one.

We can focus on a specific domain to sophisticate a sentence according to a target audience. Given a controlled vocabulary, our method is able to select the most suitable words that match a context level (see Section 3).

[1] `http://lcl.uniroma1.it/babelnet/`
[2] `http://wordnet.princeton.edu/`
[3] `http://words.bighugelabs.com`

2.4 Sentence Checker

To validate the lexical replacements made in the previous step, we are aided by the inherent knowledge of search engines. As search engines crawl content created by humans on the Web, the co-occurrence of words in the same sentence implicitly represent the common sense. Thus, the validation process of the new sentence generated begins by splitting the sentence into chunks of a predetermined size (i.e. windows size of a sentence) and querying on search engines to check if these chunks occur on the Web in high scale. If they occur, then the new sentence is validated. Otherwise, no changes are made on the sentence.

3 Evaluation

Our evaluation aims at validating the methods with respect to preservation of the original meaning and its grammatical correctness. Basing on native English speakers, our main goal is to validate our complicating process regarding potential errors introduced by our method and to check if the texts preserve their original meaning. Thus, in this evaluation, we present to the participant a text retrieved from our dataset as well as its complicated form. The questionnaire for the native English speakers is composed by the following questions: (1) Do the texts above have the same meaning? (yes/no); (2) Is the text free from grammar errors? (yes/no).

As for the dataset, we used in total 1325 sentences pairs extracted from the Terence corpus [1]. For each book in the Terence corpus, we tokenized the sentences using the Stanford NLP tool to keep the sentence structure.

The complicator tool contains many parameters for each of which the settings must be specified. Here, we describe the parameters for setting the synonym source, the controlled vocabulary and the windows size of the sentence checker. Our goal is to provide a tool that can be adapted to a specific context.

Synonym source: This parameter is used to control the synonyms suggested for a given word. In our experiments we used WordNet and BigHugeLabs (described in 2.2).

Controlled vocabulary: This parameter is used to customize the simplification to a target audience. Although the list of synonyms provides words with the same sense, a specific word might not be used by a target audience, thus the controlled vocabulary will assist in picking up the right synonym in a given context. We used four vocabularies, (i) Age 7-9 Level, (ii) Age 9-11 Level, (iii) Age 9-11 Level and (iv) Search Engine.

Window sizes: This parameter defines the boundaries of a sentence. The set of words will be checked regarding its popularity, i.e., to prevent obscure and rare sentence formulations. We set the window size between 1 and 3.

4 Results

The questionnaire was answered by 30 native English speakers and covered all sentences in the dataset (original and complicated sentences).

Table 1 presents the results of the evaluations with native English speakers. The column ‘Complicated sentences’ shows the percentage of sentences that were, to some extent, modified by the methods. The column 'Precision (same meaning)' shows the agreement of the evaluators regarding the sense similarity between the original and the simplified sentence; the column 'Precision (grammatically correct)' shows the rate of the sentences that were simplified and were free from grammatical errors.

The results are also discriminated regarding their different configuration settings for which we vary the window size, the controlled vocabulary and the synonyms source.

Table 1. Results of the complicator method for different strategies (parameter settings) from the evaluation with native English speakers

Strategy ID	Window's size	Vocabulary source	Synonym source	Complicated Sentences (%)	Precision (meaning) (%)	Precision (grammar) (%)
S_1	1	Age 7-9 Level 4	WordNet	19.92	68.32	79.21
S_2	1	Age 7-9 Level 4	BigHugeLabs	75.2	67.28	59.42
S_3	2	Age 7-9 Level 4	WordNet	3.32	81.25	81.25
S_4	3	Age 7-9 Level 4	BigHugeLabs	38.48	67.86	67.35
S_5	1	Age 9-11 Level 1	WordNet	5.27	69.23	69.23
S_6	1	Age 9-11 Level 1	BigHugeLabs	59.38	62.05	55.45
S_7	2	Age 9-11 Level 1	WordNet	1.37	83.33	66.67
S_8	3	Age 9-11 Level 1	BigHugeLabs	15.82	65.00	68.75
S_9	1	Age 9-11 Level 4	WordNet	10.74	64.81	72.22
S_{10}	1	Age 9-11 Level 4	BigHugeLabs	0.2	0	0
S_{11}	2	Age 9-11 Level 4	WordNet	2.54	91.67	91.67
S_{12}	3	Age 9-11 Level 4	WordNet	2.34	72.73	90.91
S_{13}	2	Search Engine	WordNet	6.45	65.63	75.00
S_{14}	3	Search Engine	WordNet	7.23	58.33	80.56
S_{15}	2	Search Engine	BigHugeLabs	42.77	68.52	63.43
S_{16}	3	Search Engine	BigHugeLabs	7.62	60.53	55.26
S_{17}	2	Age 9-11 Level 1	BigHugeLabs	0.2	0	0
S_{18}	3	Age 9-11 Level 1	WordNet	0.59	100.00	100.00
S_{19}	2	Age 9-11 Level 4	BigHugeLabs	13.48	69.12	61.76
S_{20}	3	Age 9-11 Level 4	BigHugeLabs	5.86	58.62	68.97
S_{21}	1	Search Engine	WordNet	58.4	77.78	80.56
S_{22}	1	Search Engine	BigHugeLabs	0.2	0	0
S_{23}	2	Age 9-11 Level 4	BigHugeLabs	0.2	0	0
S_{24}	3	Age 9-11 Level 4	WordNet	3.32	75.00	87.50

5 Discussions and Conclusions

The results show that strategies S_2 (Age 7-9 Level 4), S_6 (Age 9-11 Level 1) and S_{21}(Search Engine) achieve the highest degree of lexical replacements. These are strongly related to the size of contextualized dictionary built for each level. The most important is the result in terms of precision, regarding meaning and grammatical correctness. For this case, we see that most of the values are above 60.0%. In fact, the overall precision of the complicator aggregating the variables (window size, vocabulary and synonym) is 66.75% for meaning and 64.76% for grammar. This rather high precision numbers support the utility and applicability of our proposed method. Additionally, we believe that the complicator can significantly improve if it is used in combination with better synonyms sources. The freely online available sources used in these experiments are overwhelmed with out-of-context synonyms.

As aforementioned, the 'complicator' supports strategies for expanding vocabulary necessary to convey ideas in a different language, social contexts or environments that might require different language skills. However, in some cases, the synonymy presented may not be suitable and, therefore, every word replaced must be evaluated by a user or group of users that will use the tool.

The dynamics generated by this substitution of words resemble the use of the dictionary and, it helps to expand the vocabulary and learn the different meanings of words and expressions. As studied by Krieger [2], the use of the dictionary can be used for development at different levels of reading and textual production, therefore it plays a key role as a didactic method for expansion and improvement of knowledge of a lexicon language.

Therefore, we believe that similarly to the dictionary the 'complicator' is able to contribute as a didactic resource in the development of skills related to the domain of a language. However, it is worth noting that to achieve the stated objectives, it is essential that the use of the 'complicator' is accompanied by a teacher or someone who has prior knowledge of this tool and that recognizes its didactic potential.

Acknowledgment. This work was supported by the TERENCE project, funded by the EC through FP7 for RTD, ICT-2009.4.2.

References

1. Arfé, B., Jane, O., Pianta, E., Alrifai, M.: Story simplification: User guide. Technical report d2.2, terence project, 2011, Technical report (2011), `http://terenceproject.eu`
2. da Graa Krieger, M.: Dicionários para o ensino de lngua materna: princpios e critérios de escolha. Revista Língua e Literatura Frederico Westphalen 6 e 7, 101–112 (2004/2005)
3. Navigli, R., Ponzetto, S.P.: Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. Artificial Intelligence 193, 217–250 (2012)
4. Nunes, B.P., Kawase, R., Siehndel, P., Casanova, M.A., Dietze, S.: As simple as it gets - a sentence simplifier for different learning levels and contexts. In: ICALT (to appear, 2013)
5. Toutanova, K., Klein, D., Manning, C.D., Singer, Y.: Feature-rich part-of-speech tagging with a cyclic dependency network. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, vol. 1, pp. 173–180. Association for Computational Linguistics, Stroudsburg (2003)

Creating Awareness of Emergency Departments Healthcare Values Using a Serious Game

Claudia Ribeiro, João Pereira, and José Borbinha

INESC-ID, Rua Alvel Redol 9, Lisbon, Portugal
Department of Information Systems and Computer Science, IST/UTL, Technical University of Lisbon, Portugal
{claudia.sofia.ribeiro,joao.madeiras.pereira,jlb}@ist.utl.pt

Abstract. The world in which medicine and healthcare institutions are managed is rapidly changing in complex and unpredictable ways. In periods of rapid change, highly adaptive organizations have competitive advantage. Therefore, training a modern, adaptive and high performing team is one of the keys to success. There is a growing body of evidence that Game-Based Learning can be highly successful in driving business results, and a variety of drivers are making it harder and harder to ignore as a candidate medium where deep and immersive learning needs to be delivered as is the case of healthcare management. In this paper we describe the Serious Game ImPROVE for training medical workers on the impacts of patient care quality and costs of different configurations of Emergency Department (ED) business processes.

Keywords: Serious Games, Simulation, Healthcare, Business Process Modelling and Simulation.

1 Introduction

Across the board - from physicians and hospitals to health plans to pharmaceutical companies - healthcare and life-science organizations are facing challenging times. Costs continue to increase, regulations keep growing, and consumers are demanding more control over their healthcare decisions. Business Process Modelling (BPM) and Business Process Simulation (BPS) are a potential solution to these problems. BPM and BPS is an important part of understanding and restructuring the activities and information a typical enterprise uses to achieve its business goals. Modelling and simulation are tools and methods that are widely used in Enterprise Engineering/Organizational study where these are considered effective, efficient and economical for organizational analysis and design [1].

Recently, attention has been given to the use of Serious Games as they provide three main benefits in promoting organizational learning, namely: (i) to orient and train new employees; (ii) to select current managers or future managers; and (iii) for ongoing management training [4]. The most important advantages of applying games in a business context, are the immediate feedback, active participation of students, learning from experience, observation of key factors

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 502–507, 2013.

in an on-the-job situation, preparation for the uncertainty of business, and the high motivation to learn created by the competitive environment [7,14,6].

In this paper we describe the Serious Game ImPROVE. ImPROVE is a 3D Serious Game based on the implementation of the Manchester Triage System [8] of a Portuguese ED hospital. ImPROVE provides a player with the ability to model the business process underlying the hospital ED and check its impacts on healthcare values using the simulation features. The main goal of ImPROVE is to assist and enhance BPM and BPS activities in order to provide two main benefits, promote organizational learning and organizational change in healthcare contexts.

2 Serious Games and BPM

Serious Games can be seen as an important response from the education technologist to these "digital natives", since they can introduce clear advantages in supporting complex learning processes and knowledge transfer [12]. When compared to traditional methods, games have the advantage of allowing the simulation of complex environments, enabling learners to experience situations that are impossible to simulate using the traditional methods for reasons of safety, cost and time [3].

Although there are a wide range of process modelling grammars, there is still a lack of connection between representation and simulation, which hinders the processes analysis and optimization [11]. Generally these grammars are composed by two dimension representation making use of shapes like circles, squares, diamonds and rectangles. Authoring such schemes is generally supported by process modeling tool suites that provide a graphical model editor [9]. More recent approaches propose to complement the model editor with a limited support to simulation, bug detection, reporting and analysis [10].

Several authors argue that these approaches experience several limitations with respect to the modelling process. They typically offer a low degree of collaboration between users, the simulations are simplistic and do not take into account the possible interaction of users during the simulation process and the interfaces are many time complex and abstract requiring an extensive training in order to use these tools efficiently [13]. Serious Games are mentioned as a promising approach for teaching and simulating BPM since due to their characteristics, they may solve or minimize many of those limitations.

More than merely providing a modulation and simulation tool, Serious Games can explore the advantages of three dimensional modelling grammars [2], placing the user as an active decision maker, thus differing from the traditional tools. Due to their collaboration capabilities, Serious Games may also enhance the process of distance collaborative process modelling allowing both analyst and domain experts to collaborate in the modelling process [5].

3 ImPROVE

ImPROVE is a Serious Game based on a real-world healthcare scenario, specifically the implementation of the Manchester Triage System (Manchester Triage Group, 2005) of a Portuguese ED hospital. In ImPROVE the player assumes the role of a department manager, responsible for restructuring the ED of a virtual hospital which can be accomplish through changing and adapting the several business processes inherent to the department functioning. The following sections will detail the game architecture and implementation, the Time-Driven Activity-Based Costing (TDABC) simulator and authoring process and finally the game mechanics.

3.1 Architecture

ImPROVE architecture, illustrated in figure 1, is divided into two main modules, the Serious Games component and the TDABC simulator. The Serious Games component is generically responsible for managing the game logic, graphic assets, interfaces and players interactions. It's based on Unity3D[1] game engine and follows a script based architecture, that encapsulates functionality as C# scripts.

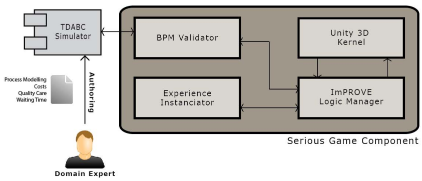


Fig. 1. Implementation Architecture

The BPM validator is a central module to ImPROVE because it allows the validation of business processes and also to retrieve information about performance (costs, patient care quality and waiting time), which can be subsequently used by the ImPROVE logic manager.

The TDABC simulator is an external module used to configure, simulate and validate the businesses processes in ImPROVE. This simulator is a dynamic linked library that can be configured using XML files, allowing a certain degree of parametrization and authoring, as well as to represent the differences between different hospitals. More details about this module are presented in the next section.

[1] www.unity3d.com

3.2 TDABC Simulator

TDABC simulator is a stand alone library, implemented in C# that provides a configurable simulation module of the Manchester Triage System. The development of this module started with the business process identification conducted together with a team of healthcare professionals through several interviews held in the hospital facilities. This process allowed us to identify key variables in this process namely: cost, patient care quality and waiting time as well as different process flows with different impacts on the identified variables. The different flows were defined by healthcare professionals and are based on small adaptations of the main process. These adaptations represent different impacts on costs, patient care quality and waiting times.

Specifically, the goals of having these variations was two fold. On one hand it was important to clearly show that every healthcare worker has a fundamental role in the overall process. This was accomplished by showing the impacts of the existence or non-existence of certain activities both for nurses regarding medical activities and vice-versa. On the other hand it was important to create awareness regarding the costs of such a complex process.

Although we integrated five different scenarios, new ones can be dynamically added to the TDABC simulator using the XML configuration. Furthermore several activity attributes can be added or changes such as: total unit time (time consumed by each unity performing a specific activity); rate (price per minute) and resources (define the healthcare professionals involved on a specific activity).

3.3 Game Mechanics

In ImPROVE the player assumes the role of a ED department manager, responsible for building and managing all the business processes underlying the hospital ED. By modelling theses processes in different ways the player is informed of its impacts on healthcare values and hospital costs using the simulation features. The game takes place on a virtual hospital ED environment in which key departments and stakeholders are present such as: patients, doctors, nurses, management staff and others, as illustrated in figure 3.3 (a). The player will be able to manage the different business process in his office using an authoring business process tool. This tool is a three-dimensional environment where the creation of business process is possible, namely using activities, decision points, start and finish primitives. The main goal of ImPROVE is to assist and enhance BPM and BPS activities in order to provide two main benefits, promote organizational learning and organizational change. In this sense, the set of activities that can be use are pre-defined and represent the range of possible activities involved or otherwise done by people in the real-world setting. Therefore, when the player is building a business process feedback is provided in order to guide the player and also to prevent designing impossible sequence of activities and or decision points. Providing real-time feedback while modelling and or simulating a business process represents an important improvement comparing to current tools. This not also motivates the player as it provides important information for decision-making.

(a) Virtual Hospital (b) Best Triage business process

The player has the ability to set up each activity duration and cost driver in real-time by accessing a context GUI. The relationships between business process primitives are manually created by the player by pressing a specific swim lane element and choosing the appropriate link. This tool is closely connected with the BPM validator module used to validate and simulate the process being modelled. An example of the ideal business process within the game can be depicted in figure 3.3 (b).

Once the business process is finished the player has the ability to test its impact by simulating it. The simulation will applied the current business process to the virtual ED department, in which the impacts can be seen using different types of feedback. The first is the feedback provided by the game HUD, displaying information regarding patient health, waiting time and associated costs. In addition the player can see the reaction of the several avatars present in the hospital, for example if the waiting time is to long, people will start to accumulate in the waiting room, if the quality of care decreases patients will decease and so on.

4 Conclusions

In this paper we describe the Serious Game ImPROVE which combines BPM and BPS for healthcare training. ImPROVE allows the player to model a business process and visually receive real-time feedback of the impacts of implementing that particular business process in a business context, which in the presented example was the ED of a Portuguese hospital. The ability to receive real-time feedback as well as visually witness the impacts of making certain decision (e.g. deciding on a particular sequence of activities or using certain resources) brings the activity of modelling business processes closer to the real world therefore, to the on-job situations. This also promotes the transformation of tacit knowledge into explicit knowledge representing a shorten period between learning a competence and being able to apply it on a concrete or similar situation. Finally, although several advantages have been pointed out, there is still a lot of space for improvements, namely adding multi-player support and collaborative tools could greatly increase knowledge transfer through socialization and promote collaboration between employees.

Acknowledgments. This work was supported by FCT (INESC-ID multiannual funding) under the project PEst-OE/EEI/LA0021/2013. The authors also would like to acknowledge to the European funded Project Games and Learning Alliance (FP7 258169) the Network of Excellence (NoE) on Serious Games.

References

1. Barjis, J., Verbraeck, A.: The Relevance of Modeling and Simu- lation in Enterprise and Organizational Study. In: Barjis, J. (ed.) EOMAS 2010. LNBIP, vol. 63, pp. 15–26. Springer, Heidelberg (2010)
2. Brown, R.A., Recker, J.C.: Improving the Traversal of Large Hierarchical Process Repositories. In: Proceedings of the 20th Australasian Conference on Information Systems, pp. 144–153 (December 2009)
3. Corti, K.: Games-based Learning - a serious business application. Informe de PixelLearning 34(6), 1–20 (2006)
4. Faria, A., Hutchinson, D., Wellington, W.J., Gold, S.: Developments in business gaming. Simulation & Gaming 40(4), 464–487 (2009)
5. Frederiks, P.J.: Information modeling: The process and the required competencies of its participants. Data Knowl. Eng., 4–20 (2006)
6. Fu, F.-L., Su, R.-C., Yu, S.-C.: Egameflow: A scale to measure learners' enjoyment of e-learning games. Compututer and Education 52, 101–112 (2009)
7. Gilgeous, V.: A Study of Business and Management Games. Management Development Review, 32–39 (1996)
8. Machester Triage Group: Emergency Triage, 2nd edn. BMJ Books (2005)
9. Magic Quadrant - Business Process Management Suites, `http://wwwimages.adobe.com/www.adobe.com/content/dam/Adobe/en/enterprise/pdfs/magic-quadrant-for-business-process-management-suites.pdf`
10. Recker, J.C.: How Good is BPMN Really? Insights from Theory and Practice. In: Proceedings of the 14th European Conference on Information Systems, pp. 1582–1593 (2006)
11. Rosemann, M., Recker, J., Indulska, M., Green, P.: A Study of the Evolution of the Representational Capabilities of Process Modeling Grammars. In: Martinez, F.H., Pohl, K. (eds.) CAiSE 2006. LNCS, vol. 4001, pp. 447–461. Springer, Heidelberg (2006)
12. Squire, K.: From Users to Designers Building a Self-Organizing Game-Based Learning. Technology Trends, 32–42 (2005)
13. West, S.B.: Collaborative business process modelling using 3D virtual environments. In: Proceeding of 16th Americas Conference on Information Systems, pp. 1–11 (2010)
14. Zantow, K.K.: More than Fun and Games: Reconsidering the Virtues of Strategic Management Simulations. Academy of Management Learning & Education, 451–458 (2005)

Serious Games Adoption in Organizations – An Exploratory Analysis

Johann C.K.H. Riedel[1], Yanan Feng[1], and Aida Azadegan[2]

[1] Nottingham University Business School, Jubilee Campus, Wollaton Road, Nottingham, UK
{johann.riedel,yanan.feng}@nottingham.ac.uk
[2] School of Computing, University of the West of Scotland, Paisley, UK
aida.azadegan@gmail.com

Abstract. This paper arises from work ongoing in the GALA (Games and Learning Alliance – Network of Excellence for Serious Games). An exploratory set of case studies were carried out to understand the benefits, barriers and enablers of adopting serious games in companies and non-educational organizations. Serious games are games that educate, train and inform. It could therefore be expected that serious games would play an important role within corporate training, but this seems not to be the case. Five exploratory case studies of SG adoption were collected. There was use of serious games for training (three cases) and for corporate change interventions (two cases). Most of the organizations commissioned the SG from an external party and only in one case did the organization itself develop the serious game. The key finding was that senior management support was critical for serious game adoption in every case. SG adoption was typically limited to a single department or branch/subsidiary of the company.

Keywords: Serious Games, Adoption, Integration, Training, Business, Industry.

1 Introduction

It could be expected that serious games would play an important role within corporate training, but this seems not to be the case. Less investigation has been focused on the integration of Serious Games (SG) in companies. The authors have been carrying out research on the adoption and integration of serious games in companies. This work is being carried out as part of the GaLA - "Games and Learning Alliance", Network of Excellence on Serious Games. This paper addresses the question of why and how serious games can be adopted/integrated into companies.

There are several ways serious games can be integrated into companies – a classification framework of the ways serious games can be integrated in to companies was devised and validated through a survey of SG cases relevant to business/industry [1]. The four ways of integrating SGs in companies are: as training, as corporate change/strategy interventions, through viral diffusion and as gamification. IT based techniques, such as Serious Games, need to be integrated in companies, however,

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 508–513, 2013.

integration is a complex and challenging process [7,11]. The literature review showed that there is a dearth of studies of serious games integration in companies. We, therefore, needed to construct a hybrid research framework based on a review of the literature on technology adoption. This framework combined the insights from the technology acceptance model [8] and the diffusion of innovation literatures [12]. The factors which can influence SG adoption were identified as: technological, organizational, environmental, contextual and individual [3,5,10] and SG specific. Five exploratory case studies of SG integration were collected from large organizations. The case studies explored the benefits of serious games, the barriers to adoption and success factors [4]. The adoption case studies have been designed as exploratory case studies. Exploratory, qualitative case studies are more appropriate to support theory development [9], and to capture and understand the social and organizational context in which the phenomenon under investigation occurs [13]. Further, when a phenomenon is still under development (as SGs are), a flexible and exploratory research design is required [6]. A qualitative research design allows both the flexibility and exploration that are required for our investigation of SG integration in organisations.

2 Analysis

The following table provides an overview of the general characteristics of the case study organisations. Most of the case organizations using serious games are large, with employee numbers above 1,000. There is a diversity of industry sectors represented, health, telecoms, banking and the military. Three of the organizations are early adopters of technology and two are first followers. This is in accord with the UK SG adoption survey [2], where 40% of respondents were either early adopters or first followers of new technology. So it seems that a positive attitude to new technology is a prerequisite for serious games use.

Table 1. Summary of Cases, SG Characteristics

	Lego Serious Play	Wallbreakers	Service Game	Afghanistan	OpenSea
SG Use Type	Corporate change intervention	Corporate change intervention	Training/ Viral diffusion	Training	Training
SGs used	Single	Single	Multiple - 3	Single	Multiple - Several
User org is SG developer	No	No	No	Yes	Yes
Commissioned SG	As consultancy	As consultancy	Yes	Internal	Internal
Deployment level	Limited (1 branch)	Limited (1 department)	Limited (1 branch)	Limited	Wide
Participant numbers	200	400	6000	500 annually	20-30

Only two types of use of serious games were present in the cases (Table 1) – for training and for corporate change intervention. One case (Service Game) used the innovative technique of viral diffusion – a distribution method of making the game available to anyone who wants to play it (and not integrating the game into formal training programmes). In most of the cases a single serious game was in use. Only in two cases was more than one serious game in use – three commissioned games in the Service Game and multiple use scenarios in the OpenSea case. In the OpenSea case the organization built a simulator which allowed experimentation with different scenarios. Most of the organizations commissioned the SG from an external party and only in one case did the organization itself develop the SG (OpenSea). In the two corporate change cases the serious game was commissioned as consultancy through the facilitator. In terms of the deployment, or internal diffusion, of serious games it was mostly limited to company branches or departments; only in the case of NATO's OpenSea was there widespread use. The numbers of participants was generally quite high – in the hundreds, except for OpenSea which was designed for low numbers of users – the emphasis being on quality not quantity.

Table 2 presents the benefits and barriers of SG use in the companies. A number of benefits of serious games were identified in the cases: serious games can get everyone involved; they are the modern way to learn and produce learning results quickly. In the corporate change cases the use of serious games signals a change and this helps to promote the cultural change objectives of the intervention. In the military cases the most important benefit was that SGs allow trainees to experience situations that would be impossible in the real world for reasons of personal safety, cost or time; they thus provide a risk-free training means. SGs were also found to be cost effective and allowed many people to be trained at the same time.

Table 2. Summary of Cases - SG Benefits & Barriers

	Lego Serious Play	**Wallbreakers**	**Service Game**	**Afghanistan**	**OpenSea**
Benefits	SGs can get everyone involved	The change in means (to SGs) signals a change in culture as well	- modern way of learning - quick to learn	- trainees can experience situations that are impossible in the real world for reasons of safety, cost or time;	- Risk-free; - train many people at the same time - train on hard to replicate situations - cost effective - quicker learning
Barriers	SG duration (1 day) too long	technical oriented company culture	high cost	- minimum computer abilities; - SG development costs	- SGs as a competitor to established programmes - Strategic change: resources, technology & space

Barriers to adoption were found to be SG duration, cost, and that minimum computer skills are required of the participants. The company culture can also be a barrier, whether being technically oriented or anti-game. Barriers during the implementation of serious games also occurred – they are competitors to existing programmes, they can require a strategic change including the reallocation of resources, the provision of PCs and dedicated physical space.

In terms of success factors the cases showed that senior management support was critical for adoption in every case. In two of the cases an SG 'product' champion was important for successful adoption. There were a number of factors which enabled the SGs to be successful in use: thorough advanced preparation, use of a facilitator, the SG being based on existing management theories, including the SG in online course platforms and the commitment of several stakeholders in the organization. For a company to develop a serious game a high level IT knowledge, especially that relevant to SGs, was necessary.

3 Conclusion

Serious games have been rising in popularity and very many games have been developed for diverse applications. However, their adoption in companies is still limited. Further, research on how to integrate serious games in companies is in its infancy. The best we can do is to rely on the general technology adoption literature and especially the IT adoption literature. This paper has sought to make a contribution to understanding how serious games can be integrated into companies, what are the benefits, barriers and success factors. A review of the literature on technology adoption was carried out to produce a hybrid research framework which was used to inform the research. The factors which can influence SG adoption were identified as: technological, organizational, environmental, contextual, individual and SG specific.

Five exploratory case studies of serious games integration in companies were carried out. They produced some insights and issues for further investigation. There was use of serious games for training (three cases) and for corporate change interventions (two cases); but not for gamification. In most of the cases a single serious games was in use. Most of the organizations commissioned the SG from an external party and only in one case did the organization itself develop the serious game. The key finding was that senior management support was critical for serious game adoption in every case. Therefore, the question of how the acceptance of senior management was gained needs to be investigated. How was senior management convinced of the benefits and by whom? Further research is required to understand the integration process – the actors involved and how the barriers were overcome. We need to develop a model of the adoption process from initial contact through generating wider buy-in, to the decision to adopt, and to implementation. And after the adoption decision who are the involved actors. Is the role of an SG champion necessary to the success of the adoption process? How were the adoption barriers overcome?

There is the question as to why the extent of SG adoption was limited; in the case studies, adoption was typically limited to a single department or branch/subsidiary of the company. So what are the factors that determine the extent of adoption, is it something to do with the customized nature, or lack of generality, of serious games?

There are a number of limitations for the current research. The first limitation is associated with the exploratory nature of the study and the limited number of cases considered. Although our purpose in using this approach was to collect reasonably good quality data which can be used to identify the important research questions for further investigations. The cases came from large companies and so we do not know if the results apply to small and medium sized companies. The number of people interviewed for the case studies was limited and we need to understand the dynamic of which persons were involved, when and why. We will then be able to identify what their motivation is and how senior management is convinced to make the adoption decision.

The case studies provided some insight into how and why serious games are integrated into companies. The next stage of the work will be to develop a model of the integration/ adoption process based on ones which have been developed for IT systems adoption. From this we can identify who the key actors could be and develop interview questionnaires for them. We can then interview the different actors involved to understand their perspectives and motivations. The extent of involvement of the HR department remains unclear. In two of the cases, the consultancy/intervention cases, HR does not appear to have been involved in the integration process; only in the Service Game case was HR involved.

Understanding serious game's value and researching into how they should be integrated to satisfy both users (companies, employees) and technology (Serious Games) providers, is part of the GALA serious games network mandate and will direct the research in the future.

Acknowledgements. The research reported in this paper has been partially supported by the European Union, particularly through the project GaLA: The European Network of Excellence on Serious Games (FP7-ICT-2009-258169) www.galanoe.eu.

References

1. Azadegan, A., Riedel, J.C.K.H.: Serious Games Integration in Companies: A Research and Application Framework. In: ICALT2012 - International Conference on Advanced Learning Technologies and Technology-enhanced Learning, July 4-6. IEEE Computer Society & IEEE TC on Learning Technology, Rome (2012a), doi:10.1109/ICALT.2012.236, ISBN: 978-0-7695-4702-2/12
2. Azadegan, A., Riedel, J.C.K.H., Baalsrud Hauge, J.: Serious Games Adoption in Corporate Training. In: Ma, M., Oliveira, M.F., Baalsrud Hauge, J., Duin, H., Thoben, K.-D. (eds.) SGDA 2012. LNCS, vol. 7528, pp. 74–85. Springer, Heidelberg (2012b)
3. Brown, I., Russell, J.: Radio frequency identification technology: An exploratory study on adoption in the South African retail sector. International Journal of Information Management 27(4), 250–265 (2007)

4. Bunduchin, R., Weisshaar, C., Smart, A.: Mapping the benefits and costs associated with process innovation: The case of RFID adoption. Technovation 31, 505–521 (2011)
5. Chan, S.C.H., Ngai, E.W.T.: A qualitative study of information technology adoption how ten organizations adopted web-based training. Information Systems Journal 17(3), 289–315 (2007)
6. Creswell, J.W.: Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Sage Publications, London (2007)
7. Damanpour, F., Wischnevsky, J.D.: Research on organizational innovation: Distinguishing innovation-generating from innovation-adopting organizations. Journal of Engineering and Technology Management 23, 269–291 (2006)
8. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 318–340 (1989)
9. Eisenhardt, K.M.: Building Theories from Case Study Research. The Academy of Management Review 14(4), 532–550 (1989)
10. Hameed, M.A., Counsell, S., Swift, S.: A meta-analysis of relationships between organizational characteristics and IT innovation adoption in organizations. Information and Management 49(5), 218–232 (2012)
11. Online Educa: Learning Technology Adoption in European Businesses, `http://www.icwe.net/oeb_special/OEB_Newsportal/learning-technology-adoption-in-european-businesses/` (last accessed: November 29, 2011)
12. Rogers, E.: Diffusion of Innovation. The Free Press, New York (1985)
13. Yin, R.: Case Study Research: Design and Methods, 4th edn. Sage, London (2009)

LIM App: Reflecting on Audience Feedback for Improving Presentation Skills

Verónica Rivera-Pelayo, Emanuel Lacić, Valentin Zacharias, and Rudi Studer

FZI Research Center for Information Technology
Haid-und-Neu-Str. 10-14, 76131 Karlsruhe, Germany

Abstract. In order to successfully give a lecture or do a presentation in a conference, presenters need certain skills as well as previous preparation. In such scenarios, reflective learning offers a great potential to improve professional skills and presenter's performance by relying on data captured during the presentation. For this purpose, we developed the Live Interest Meter (LIM App) which allows capturing, aggregating and visualizing live feedback from the audience. After developing the first prototype, testing it and conducting a user study, we developed the second prototype presented in this paper. This further development made emphasis on the recalling and revisiting of past experiences by exploring the collected data. We conducted the evaluation of the LIM App with three university lectures. Our evaluation showed positive results regarding the capturing and sharing of feedback to improve presentation skills. Whilst the LIM App guided the lecturers to reflect and to remember their presentations better, some time and advice to get accustomed to using it is still needed so that it is optimally integrated in their presentations.

Keywords: Reflective Learning, Live feedback, Tool, Quantified Self.

1 Introduction

Lectures and conferences are main daily activities in several professional fields, like teaching, research or business. The presenter needs certain skills as well as previous preparation in order to successfully address her audience. Additionally, it becomes more challenging when the size of the audience increases. In such scenarios, reflective learning, i.e., learning from own experiences, offers a great potential to improve professional skills and presenter's performance by relying on data captured during the lectures or presentations [8].

There are several approaches related to capturing data in education (e.g. [4,5,6,9]) with the purpose of increasing the interaction with the audience, but none of them focus on using the captured data to review the presenter's performance and improve by reflecting on it. Presenters can only perceive their performance from their own perspective, and having information about the perspective of the audience may offer potential for learning and improvement. This audience's perspective can be captured in form of feedback about the presentation (Is the lecture difficult or easy to follow? Is the presentation too fast or too slow?).

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 514–519, 2013.

To address these issues, we developed the Live Interest Meter App (LIM App) [7], which allows to track and visualize feedback from the audience to support reflective learning. Based on the conclusion of our previous work, this new version improves existing features and adds new ones, especially for the support of recalling and revisiting past experiences. Following the model described in [8], the LIM App concretizes an example of a Quantified Self (QS) application to support reflective learning, with the aim of guiding learners to improve their presentation skills and performance when addressing an audience. We present in this paper the second prototype of the LIM App and the evaluation we conducted in three university lectures.

In the following, we will present a review of related work (Section 2), the Live Interest Meter App (Section 3) and the conducted evaluation (Section 4).

2 Related Work

Audience Response Systems (ARS), also known as clickers [4,6,9], enable lecturers in a large lecture class to instantaneously collect student responses to a posted question, generally multiple choice [3]. These systems generally aim at improving student outcomes (e.g. exam scores or passing rates), student comprehension, and learning as well as student attendance and interest on the course. Among the commercial clickers we can find ShakeSpeak[1], a plug-in for MS PowerPoint that allows users to insert questions in the slides. Socrative[2] is also an online mobile student response system to engage classrooms through a series of educational exercises and games. There are also some approaches that explore all way communication through micro blogging [1,5] to support a more active learning experience in spite of the size and layout of large lecture halls.

Apart from the above mentioned clickers, there are other projects to support real-time feedback during lectures. NUKATH [2] was a project aiming to develop and assess the usage of notebooks in academia. Currently, the NUKATH tool is called nuKIT[3] and serves as a polling service during the lecture, also allowing to give feedback about the speed of the lecturer. WIL/MA Toolkit [10] was developed to facilitate interaction and a bidirectional and synchronous communication between lecturer and students. Another example is GoSoapBox[4], a commercial classroom response system that aims at enhancing students engagement. It has a confusion barometer, social Q&A to post a question at any point during the lecture, discussions that the instructor suggests, anonymised polls and quizzes.

Most of the work about clickers and feedback tools explores and focuses on giving a benefit to students, and are tightly related to the lectures' content and the knowledge they acquire, being limited to polling only during the lecture in many cases. This is associated to the fact that their main goal is to improve

[1] ShakeSpeak, `www.shakespeak.com`
[2] Socrative, `www.socrative.com`
[3] nuKIT: KIT App für Live-Abstimmung und -Feedback im Hörsaal, `http://elearning.studium.kit.edu/179.php`
[4] GoSoapBox, `www.gosoapbox.com`

the learning process of the students, while our work is focused on improving the learning of the lecturer in her professional activity. Therefore, we consider the lecturer as the center of the scenario and focus on the improvement of her presentation skills and performance with the use of the LIM App. In this case, students play a very important role, as they are the ones who provide feedback that serves for the reflection process and therefore they also have to perceive a benefit on the tool. For this purpose, the feedback meter function was enriched with other features that support the lecture, like polls and questions, and are intended to improve the lecture itself, but also to add context to the feedback given to the lecturer.

3 Live Interest Meter App

The second LIM App prototype focused on supporting the revisiting and recalling of past lectures. It was developed as a web application to facilititate the access to the captured data from many devices. The scenario that was considered to design the LIM App involves a person addressing a large audience, e.g. in a lecture or a conference. In this situation, three use cases can be defined. Firstly, a presenter can use the application to gather feedback (a) which can be evaluated and learned from during the presentation. Having the role of an audience member, the user can give feedback (b) during the presentation and, while doing that, improve own attitude, attention and concentration. Lastly, after the presentation, the presenter can evaluate (c) and reflect on the gathered data.

The core component of the LIM App is the meter, which has one dimension and two captions (up and down) and allows to send feedback to the presenter. Before starting, the presenter defines a quantifiable aspect of the presentation to be tracked (e.g., comprehension or speech speed) as well as the suitable color scheme for the meter (blended colours of traffic lights or blue gradient). The presenter can see a dashboard (see Fig. 1) displaying the meter with the current aggregated value, the number of currently active participants, a countdown timer, a notification area and a quick preview table of polls suggested by the audience. The LIM App also offers two additional features to improve their interaction and contextualize the feedback captured with the meter: anonymous questions to the presenter with collaborative voting and polls with several options. Finally, the audience can also send anonymous comments related to the feedback of the meter.

The audience dashboard shows the aggregated live group feedback in an evolution timeline, which displays the personal meter value, group's minimal, maximal and average values in any particular point in time (x-axis) during the presentation. Also displayed in the evolution graph are marked topics, key points during the presentation that can be entered by users in the audience.

To facilitate the revisiting of past experiences, the enrichment and presentation of the gathered data is needed. Users can go to the My events page and explore the reports of their presentations, including the meter feedback, polls, questions and written feedback comments (see Fig. 2).

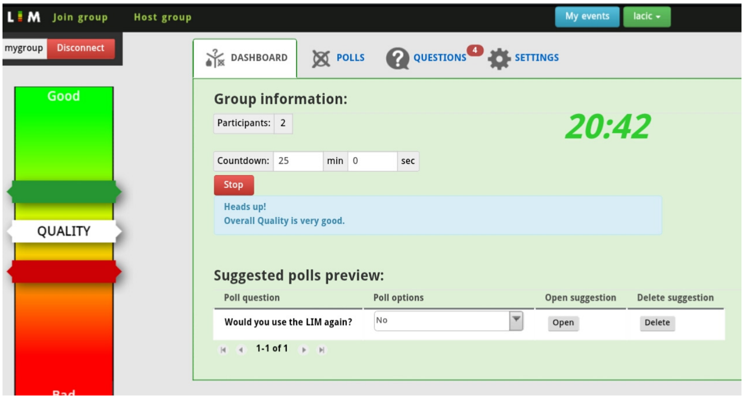


Fig. 1. Dashboard as seen by the presenter during the lecture or presentation

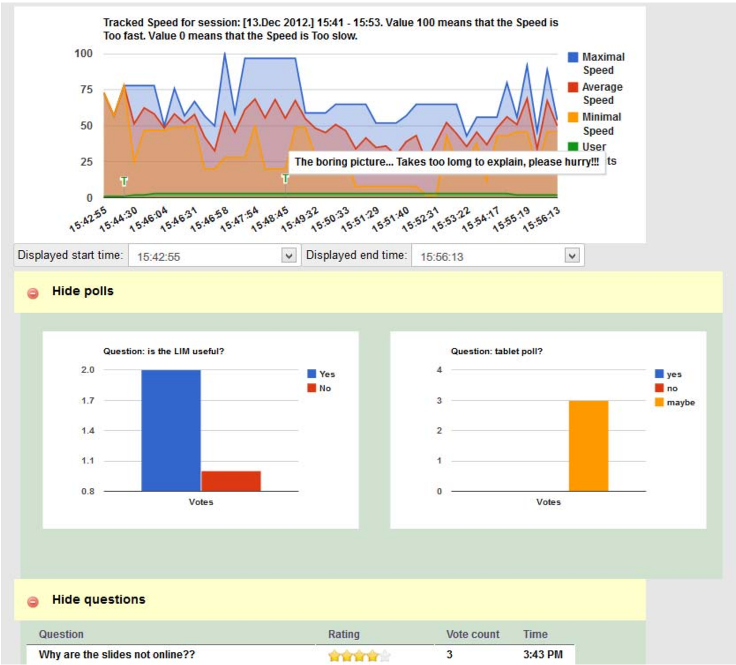


Fig. 2. Report of a presentation held by a lecturer, showing (i) the feedback timeline with the number of active participants, (ii) the results of asked polls and (iii) the posed questions with votes

4 Evaluation

The LIM App was tested and evaluated in three different lectures (see Table 1). The main lecture (Group M) was using the LIM App during three consecutive weeks. The additional two lectures used it only in one session: Group S had a highly active and small audience whereas Group L had a large audience with much less student activity.

Table 1. Description of the three lectures participating in the evaluation

Group	Sessions	Participants[5]	Degree	Field	Tracked aspects
S	1	11/13	Masters	Mathematics	Speed,
M	3	18/19	Bachelor	Computer Science	Speed, Comprehension
L	1	15/50	Bachelor	Mgmt. and Economics	Quality

The questionnaire addressed four main topics: features of the LIM App, level of participation, thoughts about learning by reflection, and outcomes and work assessment using the LIM App. One researcher of the LIM team was present at the lectures to assess the presenter and students as well as to follow the course of the evaluation. On average the LIM App was used for 83 minutes. The polling function was only used by the presenters in Group M and L. Every presenter used the question functionality at least once in their lecture and there were 17 questions in the lecture with big audience (L).

One presenter agreed strongly with the LIM App helping her to reflect on experiences from work. On the statement that the LIM App has helped by providing information relevant for the decision to reflect one presenter did not agree (Group S), but the other two did. The same two presenters agreed that the LIM App helped them to collect information relevant to reconstruct work experiences as well as to actually remember and reconstruct their lectures. They also agreed that relevant content is provided for the reflection to take place, where the presenter from Group S remained neutral. Allowing the audience members to send general feedback has proved to be a great support for reflection. Related to the main goal of using the LIM App, all three presenters agreed that it is important to improve their presentation skills. In the case of Group M, the lecturer stated that he would use it regularly as it allows the comparison of different teaching strategies. In Group L, it was not clear for the lecturer if the LIM App had distracted the students. As for Group S, the answer was negative, being affected by the high interaction already existing in this group.

In order to be motivated to give their feedback, members of the audience should also perceive a benefit. Therefore, our evaluation also delivered insights about how they could improve their learning experiences. The results from their perspective were overall positive. Some students stated that (1) the LIM App is fun, (2) better than the usual way of giving feedback, (3) provides a possibility to participate personally in the lecture, (4) allows anonymous questions and feedback and (5) with it, the lecturer is able to adapt to the needs of the students.

[5] Number of valid responses / Total number of students in the lecture.

5 Conclusion

In this paper, it was shown with the developed Live Interest Meter application (LIM App) how learning by reflection can be supported with a Quantified Self tool, which allows quantifying and tracking feedback from the audience. In accordance to the analysis of the questionnaires, reflection took place by the users who used the LIM App most during and after the presentation. A key factor in collecting data for reflection is the presenter herself and how she engages the audience to participate and how she integrates the LIM App in the lecture. In order to optimally achieve this, a brief training is needed. Taking the results into consideration, it is clear that the target area for the LIM App are indeed presentations with a large audience number, and little or no student - lecturer interaction. Future work will include a longer evaluation at larger scale.

Acknowledgements. The project "MIRROR - Reflective learning at work" is funded under the FP7 of the European Commission (project number 257617).

References

1. Akbari, M., Böhm, G., Schroeder, U.: Enabling communication and feedback in mass lectures. In: ICALT, pp. 254–258 (2010)
2. Bonn, M., Dieter, S., Schmeck, H.: Kooperationstools der Notebook Universität Karlsruhe (TH). In: David, K., Wegener, L. (hrsg.) Mobiles Lernen und Forschen, pp. 63–71 (November 2003)
3. Caldwell, J.E.: Clickers in the large classroom: Current research and Best-Practice tips. CBE Life Sci. Educ. 6(1), 9–20 (2007)
4. Duncan, D., Mazur, E.: Clickers in the Classroom: How to Enhance Science Teaching Using Classroom Response Systems. Pearson Education (2005)
5. Hadersberger, J., Pohl, A., Bry, F.: Discerning actuality in backstage - comprehensible contextual aging. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 126–139. Springer, Heidelberg (2012)
6. Kundisch, D., Herrmann, P., Whittaker, M., Beutner, M., Fels, G., Magenheim, J., Reinhardt, W., Sievers, M., Zoyke, A.: Designing a Web-Based Application to Support Peer Instruction for Very Large Groups. In: ICIS 2012, Research in Progress, Orlando, USA (December 2012)
7. Rivera-Pelayo, V., Munk, J., Zacharias, V., Braun, S.: Live interest meter: Learning from quantified feedback in mass lectures. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge, LAK 2013, pp. 23–27. ACM, New York (2013)
8. Rivera-Pelayo, V., Zacharias, V., Müller, L., Braun, S.: A framework for applying quantified self approaches to support reflective learning. In: IADIS Mlearning 2012, Berlin, Germany (2012)
9. Rubner, G.: mbclick - an electronic voting system that returns individual feedback. In: WMUTE, pp. 221–222. IEEE (2012)
10. Wessels, A., Fries, S., Horz, H., Scheele, N., Effelsberg, W.: Interactive lectures: Effective teaching and learning in lectures using wireless networks. Comput. Hum. Behav. 23(5), 2524–2537 (2007)

Applying the Widget Paradigm to Learning Design: Towards a New Level of User Adoption

Bernd Simon[1], Michael Aram[1], Frans Van Assche[2],
Luis Anido Rifon[3], David Griffiths[4], and Manuel Caeiro Rodríguez[3]

[1] Knowledge Markets Consulting, Siebensterngasse 31/12, A-1070 Wien, Austria
{bernd.simon,michael.aram}@km.co.at
[2] University of Leuven, Computer Science Department, Belgium
frans.vanassche@cs.kuleuven.be
[3] Universidad de Vigo, Campus Universitario de Vigo,
As Lagoas, Marcosende, ES-36310 Vigo
{luis.anido,mcaeiro}@det.uvigo.es
[4] University of Bolton, Deane Road, BL3 5AB Bolton, United Kingdom
dai.griffiths.1@gmail.com

Abstract. Researching the sharing of learning designs is a well-established domain within the technology-enhanced learning research community. However, until now tools supporting educational modelling languages such as IMS Learning Design have reached a wide adoption in today's school practice. Following a design science research methodology we report on the design, implementation, and evaluation of a novel tool referred to as "Composer". The Composer supports the design of learning activities and has been developed according to design principles such as (a) interoperability between design-time and run-time systems based on the W3C Widget Standard, (b) inclusion of artefact types beyond content such as tools, people and events, (c) a user-friendly authoring environment. A first evaluation of the proof-of-concept implementation suggests that the tool is easy-to-use and provides added value for teachers when it comes to reflecting about the design of learning activities.

Keywords: Learning Design, Educational Modelling, W3C Widgets, Mashups, Wookie, Design Science.

1 Motivation and Methodology

The primary role of any instructional agent such as a teacher, is to stimulate the performance of learning activities that will gradually result in the attainment of the learning outcomes [1]. As a consequence learning design has always been of a particular interest to the educational domain. [2] provide an overview of learning design authoring software that aims to simplify the learning design process. [3] provide a graphical user interface for designing learning activities based on IMS Learning Design. Evaluation of the latter revealed the disconnection between the design tool and the runtime system as one major problem with respect to user acceptance [3]. Until now,

D. Hernández -Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 520–525, 2013.

tools supporting educational modelling languages have not reached wide adoption [2], [4], [5]. In the following, learning design is meant as the preparation of a unit-of-learning (e.g. course, lesson) and includes the definition of learning outcomes, the selection of learning resources, and the sequencing of measures [1], [4].

We focus on (blended) learning environments of the school sector. Although evidence suggests that ICT can have a positive impact on the expansion of learning opportunities [6], [7], there is still a significant number of schools in Europe that lack sufficient computer equipment [7]. At the same time the adoption of ICT also varies between subjects [8].

This research effort is tackling the learning design issue by applying a design science research methodology (DSRM) [9] on the problem. Design science research is a research paradigm in which the researchers try to answer questions to relevant problems via the creation of innovative artefacts [10], [11]. A number of efforts have been made to describe and guide construction-oriented research processes [12], [13], [14]. Based on this the DSRM identifies six activities [9], i.e. problem identification and motivation, objectives for a solution, design and development, demonstration, evaluation, and communication.

This paper is structured accordingly. While Section 2 presents requirements and reflects them in the context of related work, Section 3 presents our proposed solution relying on a widget-based architecture. In Section 4 the paper concludes with preliminary findings of our evaluation.

2 Requirements and Related Work

Our proposed solution is centred on the design and facilitation of Learning Activities. We define our key concepts as follows: a *Learning Activity Design* (LAD) describes a discrete session of Learner interactions, including potential Learning Resources to be used, in order to achieve educational outcomes. We adopt a broad view of the term "*Learning Resource*" as to comprise Content, Contributors, Events and Tools and assume that all artefacts that are meant to become part of the final learning experience are represented as, or delivered through, widgets. A conceptually similar model is proposed by the Simple Learning Design 2.0 [4] specification, where a learning design is manifested in a so-called scenario, which is composed of learning activities that can also include learning objects.

Our relevant stakeholder roles include *Learners*, *Teachers*, and *Pedagogical Coordinators*. A Pedagogical Coordinator inspires other Teachers to adopt pedagogical innovation mediated by LADs in a learning context that aims to serve Learners in attaining learning outcomes.

Driven by an analysis of the educational context we have identified the following requirements – very similar to those identified by [15] for learning design repositories. Pedagogical Coordinators demand means to create and reuse LADs and share these in order to inspire Teachers. While [15] identify the need for downloading learning designs, we go further by proposing a complete architecture for the exchange of learning designs. Teachers need to easily find and assess a LAD in order to make the associated Learning Resources available to Learners. To support the exchange of LADs across systems the systems involved need to be interoperable. *Interoperability* is referred to as the ability of systems or components to use information that has been exchanged [16]. Research distinguishes between interoperability on the object, referring to a proper use of the information provided – and interoperability in communication, referring to an agreed communication protocol between systems [17]. These two aspects of interoperability translate into a requirement for making LADs re-useable in different technical contexts as well as for agreeing on communication protocols.

3 Proposed Solution: A Widget-Based Architecture

We have designed an architecture that consists of the following components: A *Widget* is a packaged web application [18] that is designed to be easily distributed and embedded within varying contexts (e.g. within a portal-style mashup, on a mobile phone, etc…). Widgets rely on open standards with respect to both their representation format and their communication protocols.

The *Widget Store* is a software component that is built on the Apache Wookie and EDUKApp technologies [19]. It supports the uploading, tagging, and searching for Learning Resources and LADs in the form of Widgets. The *Composer* supports Pedagogical Coordinators and Teachers in designing Learning Activities, and augmenting them with Learning Resources.

A *Widget Run-time Environment (RTE)* acts as the "entry point" for end users and is a configurable software container that provides an environment allowing users to identify and add their Widgets and to integrate them in order to meet the educational objectives of a Learning Activity. Typically, a Widget RTE connects to a Widget Store to provide users with an integrated experience when selecting and instantiating widgets [20]. Examples include mashup engines like Apache RAVE as well as Widget-enabled learning management systems like Moodle and DotLRN.

3.1 Representing Learning Designs via Widgets

In the following we describe our layered approach to representing LADs, that follows the "web best practices" of *progressive enhancement* and the *rule of least power* [21]. Consequently, when entering a higher level, interoperability decreases, while functionality increases. At the lowest layer we render a LAD as HTML. Hence, the fundamental (narrative) information of such a guide is represented as a web document, thus can be viewed in any standard web browser, or processed otherwise by third-party applications.

Packaging this LAD as a W3C Widget represents the second layer, which allows teachers to (easier) use the LAD in various manners, e.g. instantiate it in their Widget RTE, view it offline on a phone, or publish it in a Widget Store. At these two levels the LAD already provides added value to the teacher both when preparing the learning activity and when it takes place.

However, many useful LADs will go beyond mere textual descriptions and will require actually useful resources (e.g. Applications, Content). Hence, the technology shall support the Teacher in augmenting the (virtual) learning environment with these resources. We therefore progress further in the functional enhancement by "transforming" the instantiated LAD Widget into a mashup. Technically, to this end our approach utilizes a client side cross-context communication channel (based on pmrpc [22]) to transmit a description of the additional Learning Resources required. As we consider all resources to be delivered via Widgets, we represent the resources required by the LAD in the form of an mashup description based on the Open Mashup Description Language (OMDL)[1]. Finally, the RTE instantiates all the Widgets required for conducting the learning activity that is described by the LAD Widget.

3.2 The Composer: Authoring Environment for Widget-Based LADs

The Composer is supposed to provide Pedagogical Coordinators and Teachers with means to compose and re-use Learning Activities and augment them with Learning Resources. We interpreted this process as Widget aggregation [21]. An important use case in the context of the Composer is the "typical usage scenario for learning resources" [17], which includes the discovery, repurposing, and re-publishing of a LAD. When repurposing the LAD the Teacher aggregates Widgets. The result of this mashup activity is a personalized LAD in the form of a LAD Widget generated using Xocp[2]. For interoperability, the Composer has been implemented as a highly embeddable web application, i.e. we allow its seamless integration via the IMS Learning Tools Interoperability (LTI)[3] protocol and implemented a responsive user interface using Bootstrap[4]. Moreover, a "Composer Widget" that implements a simple wrapper mechanism makes the application also accessible via the Widget Store.

4 Evaluation

At the time of writing 19 small-scale evaluation activities of the proposed solution have been carried out. Early evaluation activities mainly consisted of open expert interviews from which a better understanding of the problem definition was derived. At a later stage these activities were used to iteratively revise the requirements. The evaluation events mostly involved pedagogical experts. Early evaluations were documented in the form of action logs resulting in concrete changes to requirements. In the case of the Composer the main findings relate to: (a) a simple user interface, (b) private areas within the collaborative, wiki-style tool, and (c) better integration with mobile devices like tablets.

---

[1] http://omdl.org/
[2] http://wiki.tcl.tk/28538
[3] http://www.imsglobal.org/lti/
[4] http://twitter.github.io/bootstrap/

Acknowledgements. This article is the result of the iTEC (Innovative Technologies for an Engaging Classroom) Project supported by the European Commission's FP7 Programme (Grant agreement Nº 257566).

References

[1] Koper, R., Bennett, S.: Learning Design: Concepts. In: Adelsberger, P.D.H.H., Kinshuk, P., Pawlowski, P.D.J.M., Sampson, P.D.G. (eds.) Handbook on Information Technologies for Education and Training, pp. 135–154. Springer, Heidelberg (2008)

[2] Derntl, M., Neumann, S., Griffiths, D., Oberhuemer, P.: The Conceptual Structure of IMS Learning Design Does Not Impede Its Use for Authoring. IEEE Transactions on Learning Technologies 5(1), 74–86 (2011)

[3] Neumann, S., Oberhuemer, P.: User Evaluation of a Graphical Modeling Tool for IMS Learning Design. In: Spaniol, M., Li, Q., Klamma, R., Lau, R.W.H. (eds.) ICWL 2009. LNCS, vol. 5686, pp. 287–296. Springer, Heidelberg (2009)

[4] Durand, G., Belliveau, L., Craig, B.: Simple Learning Design 2.0. In: IEEE 10th International Conference on Advanced Learning Technologies (ICALT) 2010, pp. 549–551 (2010)

[5] Durand, G., Downes, S.: Toward Simple Learning Design 2.0. In: 4th International Conference on Computer Science & Education, ICCSE 2009, pp. 894–897 (2009)

[6] Core ICT Indicators 2010. International Telecommunication Union (ITU)

[7] Technology, Broadband and Education Advancing the Education for All Agenda - A Report by the Broadband Commission Working Group on Education. International Telecommunication Union (ITU)

[8] PISA, Results: Students On Line Digital Technologies and Performance. OECD (2009), http://www.oecd.org/pisa/pisaproducts/pisa2009/48270093.pdf

[9] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24, 45–77 (2007)

[10] Hevner, A., Chatterjee, S.: Design Research in Information Systems: Theory and Practice. Springer (2010)

[11] March, S.T., Smith, G.F.: Design and Natural Science Research on Information Technology. Decision Support Systems 15(4), 251–266 (1995)

[12] Hevner, A.R.: A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19(2), 87 (2007)

[13] Vaishnavi, V.K., Kuechler, W.: Design Science Research Methods and Patterns: Innovating Information and Communication Technology, 1st edn. Auerbach Publications (2007)

[14] Takeda, H., Veerkamp, P., Tomiyama, T., Yoshikawa, H.: Modeling Design Processes. AI Mag. 11(4), 37–48 (1990)

[15] Sampson, D.G., Zervas, P., Sotiriou, S.: From Learning Objects Repositories to Learning Design Repositories: The COSMOS Learning Design Repository. In: 2011 11th IEEE International Conference on Advanced Learning Technologies (ICALT), pp. 285–289 (2011)

[16] IEEE Standard Computer Dictionary. A Compilation of IEEE Standard Computer Glossaries (1991)

[17] Van Assche, F., Duval, E., Massart, D., Olmedilla, D., Simon, B., Sobernig, S., Ternier, S., Wild, F.: Spinning Interoperable Applications for Teaching & Learning using the Simple Query Interface. Journal of Educational Technology & Society 9(2), 51–67 (2006)

[18] Packaged Web Apps (Widgets) - Packaging and XML Configuration (Second Edition) W3C Recommendation (2012), `http://www.w3.org/TR/widgets/`

[19] Griffiths, D., Johnson, M., Popat, K., Sharples, P., Wilson, S.: The educational affordances of widgets and application stores. Journal of Universal Computer Science 18(16), 2252–2273 (2012)

[20] Soylu, A., Mödritscher, F., Wild, F., De Causmaecker, P., Desmet, P.: Mashups by Orchestration and Widget-based Personal Environments: Key Challenges, Solution Strategies, and an Application. Program: Electronic Library and Information Systems 46(4), 383–428 (2012)

[21] Berners-Lee, T., Mendelsohn, N.: The Rule of Least Power. W3C (2006)

[22] Zuzak, I., Ivankovic, M., Budiselic, I.: Cross-context Web Browser Communication with Unified Communication Models and Context Types. In: MIPRO, 2011 Proceedings of the 34th International Convention, pp. 690–695 (2011)

Identifying Problem-Based Scaffolding Patterns in an Online Forum for Construction Professionals

Kairit Tammets, Mart Laanpere, Tobias Ley, and Kai Pata

Tallinn University, Institute of Informatics, Narva Road 25, 10120, Tallinn, Estonia
Kairit.Tammets@tlu.ee

Abstract. Online help seeking is a common form of informal learning, and it has been studied extensively in a number of different settings. In an attempt to better understand how help is provided in online help seeking for a physical work environment, we have performed an in-depth qualitative study on questions and answers found in the largest Estonian online forum for construction field. Our intention is to understand the types of learning problems identified in the help-seeking process in the online forum and how problem-oriented scaffolding is provided and received in a networked setting. We used Jonassen's typology for categorizing the similar patterns in problem-based learning that takes place in the forum. Data was analysed from two perspectives: first the problem definitions identified in forum threads were analysed based on the Jonassen's typology, along with corresponding solution paths and solutions. The same discussion threads were then analysed also from the perspective of scaffolding patterns – how people support each other and how the help is being acknowledged. The results of our analysis point out that the human supporters provide certain scaffolding patterns for supporting help-seekers to solve problems.

Keywords: Forum, Communities of Practice, Workplace learning, Scaffolding.

1 Introduction

Help seeking at the workplace is an important form of informal learning, as it enhances the zone of proximal development [1] by triggering scaffolding activities provided to a help-seeker by a more capable peer. In this paper we explore how work-related informal learning takes place in an online forum used by Estonian construction specialists and how better scaffolding could support such learning. Sawyer [3] has defined scaffolding as "*the support given during the learning process which is tailored to the needs of the student with the intention of helping the student achieve his/her learning goals*". Traditionally, construction specialists share their work-related practices in the face-to-face manner and therefore hardly any traces are left about the informal professional learning in the construction domain. The aim of this paper is to investigate what kind of help is sought and what kind of support is provided among construction professionals in the largest Estonian online forum in

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 526–531, 2013.

respective domain. For doing so, we use the typology of learning problems proposed by Jonassen [2] because we found striking similarities between Jonassen's problem types and the categories of help-seeking forum posts identified through our initial open coding exercise.

As online help seeking almost always triggers a series of (eventually fading) interactions between the help-seeker and more capable peer(s), it could be effectively interpreted as a scaffolding process. Informal learning often occurs in Web-based forums when people seek for help, share the knowledge, and learn from others' practices and experiences. In traditional scaffolding models mainly two roles are handled: tutor and tutee [4]. In the era of networked learning there is a need for the model that addresses the scaffolding phenomenon in online settings. Our study is guided by the following research questions: a) What are the main types of problem-related help-seeking patterns identified in the online forum of construction specialists?; b) What kind of scaffold the members of the construction forum provide to the help-seekers and how it is being acknowledged?

As a result the study identifies an initial set of networked scaffolding patterns. Such patterns can be used to inform design decisions when building next-generation socio-technical systems that support workplace learning in construction field. Peer scaffolding can be used in the system when helping to identify the problem type. Semantic systems may suggest then solution paths or pre-described solutions for solving the problem.

2 Theoretical Framework

We interpret the members of Estonian online forum for construction specialists as a Community of Practice (CoP): a self-organized group of individuals concerned with a specific practice, who are learning how to improve this practice through regular interaction [5]. Online communities of practice (also Internet-mediated communities of practice) [6] have the added element of "facilitative" technology, which facilitates or encourages communication. Hoadley [7] has argued that scholars have mostly focused on the role of technology in supporting the communication within CoP rather than the effects of the technology on the shared practices of community. In the construction forum mainly the communication (help-seeking, knowledge and information sharing etc.) occurs, also the practice of the construction can be supported by following the advice of the community experts.

Scaffolding was proposed as means to bridge the gap between what students can do on their own and what they can do with the help of a more capable other [8]. Technological advancements have allowed enhancing the tutor- or teacher-based scaffolding with the computer-based scaffolding of self-directed learners [9]. However, according to Saye and Brush [10] computer-based scaffolding cannot be sufficient and it needs to be accompanied by human-facilitated scaffolding. In this article the human-facilitated scaffolding with the technological support means that the members of the online CoP support each other in the problem-solving process. While initially, the use of technology in help seeking was seen as a mere mediator (moving the conversation online from face-to-face and thereby achieving a greater scope), it has recently become clear that scaffolding in networked online environments is different from traditional scaffolding. As opposed to traditional scaffolding context based on face-to-face interaction between a single tutor and tutee, in networked environments the "best" answer to a problem may emerge from iterative, aggregated and sometimes even conflicting inputs provided by several contributors. Understanding and defining the problem, finding alternatives and the best solution, then, is a negotiation process to which potentially many can contribute. Because these negotiation processes are visible for everyone in the community, it can be expected that each single interaction has an effect on the professional knowledge accumulated in the community as a whole, e.g. on how the problems are defined and categorized, who can be trusted as an expert, which solutions are preferred and what practices are promoted (both in terms of construction and in terms of how people learn construction related knowledge and practices). So the online forum can function as an amplifier of the negotiation on identity, practices and knowledge shared by the CoP.

3 Research Design and Methods

The Web-based construction forum *Ehitusfoorum* for Estonian construction professionals and amateurs was used as the main context and data source in the current study. Forum was launched in 2003, today it has more than 10 000 registered users with nearly 200 000 posts in 30 000 threads, annotated with 34 different category labels. The data from the forum was analyzed qualitatively by focusing what problem types occur in the construction forum when the members seek for help and what kind of scaffold can be identified in the help-seeking process.

A set of help-seeking discussion threads were identified and categorised using open coding method in the first phase of the study. Then categories were compared with some existing typologies of learning problems and identified a good match with Jonassen's problem-based learning typology. The next, we tested how his typology can be used for analysing problems that occur in the construction online forum. Jonassen comprehensive typology for categorizing different types of problems and their nature and characteristics are: a) logical; b) algorithmic; c) story; d) rule-using; e) decision-making; f) design problems; g) trouble-shooting; h) diagnosis-solution problems; i) strategic performance; j) case analysis; k) dilemmas with the multiple answers. We collected more than hundred discussion threads from the forum and classified them to Jonassen's problems types. Then we identified in each discussion thread the problem statement, the solution path and solution. In most of the cases it was important to keep one thread as the unit of analysis. It means that by analysing only the problem as it was formulated by the help-seeker, which was often one question, most of the problems in construction forum would be as trouble-shooting problem types. Usually the discussion between the help-seeker and helpers formed the question and it was possible to classify it to other types as well. In the next phase the same discussion threads were analysed in terms of scaffolding patterns that appeared in the problem-solving process. The final decision about classification was negotiated between two researchers; dissents were discussed until the consensus was achieved.

4 Results and Discussion

A set of 100 problem-oriented questions posted to the Web-based forum was mapped to Jonassen's problem types. It was faced that often the problem statement expressed by help-seeker is not enough for identifying the problem type. Often the problem can be classified after seeing comments, questions and solution paths provided by members. E.g. an example of the problem statement: *I would like to put the fiberboard to my floor. Is there any difference where should be the connecting link/point between the boards?* It can be decision-making type of problem when the help-seeker identifies benefits and limitations, selects alternatives and justifies. But the problem solution by the community member is following: *Logical step would be to finish the board on top of the beam.* Therefore the problem can be classified as logical, because solving the problem (connecting link between the boards) is like solving the puzzle and the boards are ended each time on the next beam.

Analysis of the discussion threads indicated that in general Jonassen's typology is suitable for categorizing the problems in construction forum, although several of his problems types (case analysis, story problems, dilemmas) are not relevant to construction field as they seems to be in educational, medicine or other contexts. Problems that occurred significantly more often were troubleshooting and diagnosis-solution problems. We assume that each of these problem types may need different type of scaffold. In this section two problem types are analyzed from the perspective of scaffolding in network. Scaffold is seen from the peer-level and networked-level interactions, which occur in the turn-taking process in online forum. Peer-level scaffold involve one or many community members who help to solve the problems of the help-seeker whereas networked-level scaffold involved accumulated community knowledge, which is stored in the system and is accessible for the other users.

Next, two problem types are analyzed from the perspective of the scaffolding patterns (Figure 1) that can be used in technological settings for supporting informal learning in the workplace: *Decision-making* problem (Figure 1) is illustrated with one of the examples from forum and that kind of problem is focusing more on comparing the possible solutions and their advantages and disadvantages. Discussions threads indicate that in case of the decision-making problems, engagement of larger group of community members in negotiation becomes more visible. Although the system-level community knowledge was also suggested, it was based on the previous shared understanding between the community members engaged in that discussion. Consensus on decisions was built through social negotiation and weighing the pros and con's of different solutions. *Diagnosis-solution problems* focus on finding out why does not system work. Figure 1 illustrates the problem type with the three examples. Although in that problem type the peer- and networked-level scaffolding was able to identify, the own experience was not so dominant. Rather it was possible to identify different types of scaffolding based on personal experience and the status or trustworthiness of the person who provides help. Barrett et al [11] have pointed participants must trust the responses they receive from the network.

Another example for networked-level scaffolding processes taking place in the forum is the accumulation of evidence over particular elements. For each problem type, different elements can be accumulated. For example, Figure 1 indicates how different solutions (in Diagnosis Problems) and alternatives (in Decision Problems) accumulate in the forum discussions. Similar examples can be found for other elements (e.g. criteria, weights of criteria, symptoms and their importance etc.).

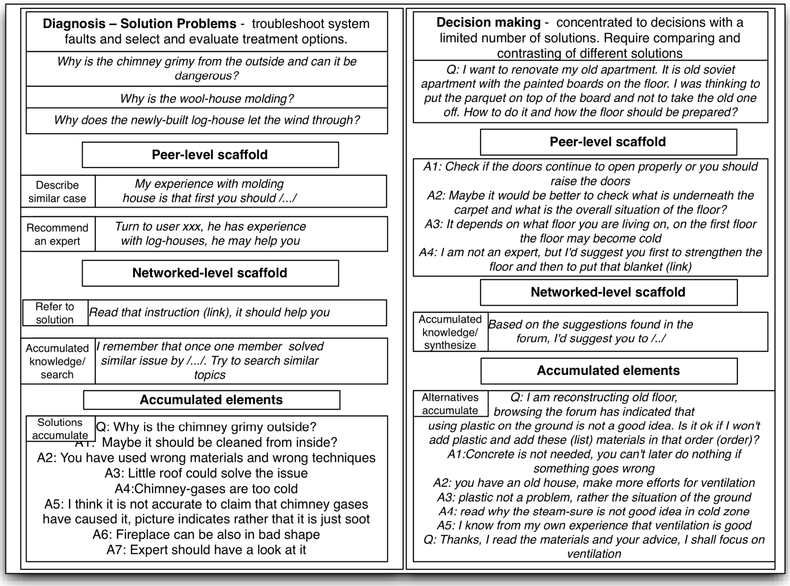


Fig. 1. Problem types and related scaffolding

Barrett et al [11] have pointed that scaffolding in the network can be provided by persons who: a) share the community identity; b) synthesize supportive knowledge from one community across community borders for aiding; c) are trusted because of their wise and reputation in the community; d) are not afraid of evaluation based on the analytics of their help-giving behaviors; e) do support voluntarily to gain peer esteem in the community; f) do not afraid of idea-stealing because they can get credit of honest attribution as the originators of the ideas. Forum analysis indicated that personal trust accumulation is important in any type of problems. Users who are trusted are highly evaluated by the help-seekers. One thing that raises the trust of the one who provides help is the functionality of peer evaluation. Under the user profile in the system, it can be seen if the user has been evaluated positively or negatively and why, for instance: *fast and good answers; taught me as the amateur with the good advice; did a good job when building my house*. Under the profile, the number of contributions made by this user, are also marked, which is used for identifying status or rank of the user such as construction veteran, construction pensioner, construction specialist. Based on peer evaluations and status, users often turn to them or suggest asking from these users (I remember that user xxx had similar experience).

Scaffolding patterns that technological system may provide for informal workplace learning include suggestions to search from the existing knowledge, visit external Web pages, ask advice from experienced users and discuss with the peers to weigh alternative solutions to the problem.

5 Summary

In this paper we took the approach that interactions in the constructor's online forum when users help for seek can be used for identifying the scaffolding patterns in the networked settings. Traditionally scaffold occurs between the tutor and tutee, but in the online CoP the concept of scaffold becomes broader. Our analysis of discussion threads in the forum indicated the support that comes from the peers provide certain scaffolding patterns for supporting help-seekers to solve problems - they try to identify together with the help-seeker the problem type and negotiate it to provide better solutions - so help provision is based on problem type classification (grounding) and using scaffolding prompts to negotiate these problem types. Additionally we point out the important role of networks in scaffolding in a way the resources are used by the members – suggested later, linked and referred to.

Acknowledgements: This study was partly funded by the Estonian Ministry of Education and Research grant 0130159s08 and 7th FP project Learning Layers.

References

1. Vygotsky, L.S.: Mind in society: The development of higher psychological processes. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (eds.). Harvard University Press, Cambridge (1978)
2. Jonassen, D.H.: Toward a Design Theory of Problem Solving. Educational Technology Research & Development (ETR&D) 48(4), 63–85 (2000)
3. Sawyer, R.K.: The Cambridge Handbook of the Learning Sciences. Cambridge University Press, New York (2006)
4. Graesser, A.C., Person, N.K., Magliano, J.P.: Collaborative dialogue patterns in naturalistic one-to-one tutoring. Applied Cognitive Psychology 9, 495–522 (1995)
5. Brown, J.S., Duguid, P.: Organizational learning and communities of practice. Toward a unified view of working, learning, and innovation. Organization Science 2(1), 40–57 (1991)
6. Stuckey, B., Smith, J.D.: Building Sustainable Communities of Practice. In: Hildreth, P., Kimble, C. (eds.) Knowledge Networks: Innovation Through Communities of Practice. Idea Group, Inc., New York (2004)
7. Hoadley, C.: What is a community of practice and how can we support it? In: Jonassen, D.H., Land, S.M. (eds.) Theoretical Foundations of Learning Environments, 2nd edn., pp. 287–300. Routledge, New York (2012)
8. Wood, D., Bruner, J., Ross, G.: The role of tutoring in problem-solving. Journal of Child Psychology and Psychiatry 17, 89–100 (1976)
9. Azevedo, R., Hadwin, A.F.: Scaffolding self-regulate learning and metacognition – implications for the design of computer-based scaffolds. Instructional Science 33, 367–379 (2005)
10. Saye, J., Brush, T.: Scaffolding critical reasoning about history and social issues in multimediasupported learning environments. Educational Technology Research and Development 50(3), 77–96 (2002)
11. Barrett, M., Cappleman, S., Shoib, G., Walsham, G.: Learning in knowledge communities: managing technology and context. European Management Journal 22(1), 1–11 (2004)

Learning, Learning Analytics, Activity Visualisation and Open Learner Model: Confusing?

Susan Bull[1], Michael Kickmeier-Rust[2], Ravi K. Vatrapu[3], Matthew D. Johnson[1], Klaus Hammermueller[4], William Byrne[1], Luis Hernandez-Munoz[1], Fabrizio Giorgini[5], and Gerhilde Meissl-Egghart[4]

[1] University of Birmingham, UK
[2] Technical University of Graz, Austria
[3] Copenhagen Business School, Denmark
[4] Talkademy, Austria
[5] Lattanzio Learning, Italy
s.bull@bham.ac.uk

Abstract. This paper draws on visualisation approaches in learning analytics, considering how classroom visualisations can come together in practice. We suggest an open learner model in situations where many tools and activity visualisations produce more visual information than can be readily interpreted.

Keywords: Learning, learning analytics, open learner model, visualizations.

1 Introduction

There is much attention on visualisation and learning analytics, based on large amounts of data for a variety of purposes [1]. *Learning* data is our focus here. Various visualisations can be used, including dashboards (see [2]). We illustrate how visualisations may be used, and unite them with an open learner model (OLM) to help teachers and students interpret the range of visual analytics that may be produced.

OLMs are concerned with visualising the learner model (individual's knowledge, competencies, etc.) OLMs commonly aim to promote metacognitive behaviours (reflection, self-monitoring, planning) [3]. So, while learning analytics often show activity data (interaction time in discussion; links in social networks or collaboration tasks; performance data), OLMs use inferences drawn from interaction to produce visualisations of the current state of the learner's understanding, competencies, and so on.

2 Next-TELL Visualisations

Next-TELL (http://www.next-tell.eu) comprises many tools and visualisations to support learning, as well as visualise activities outside Next-TELL tools, such as Moodle discussions or Google Docs. Many visualisations display similar information in a similar form as other approaches (e.g. [2]). Figure 1 shows visualisations from a Moodle discussion to illustrate the type of information and visualisation available for user interpretation. As in other learning analytics solutions, they indicate what is happening (here: interactions between learners, topics of discussion, word counts and

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 532–535, 2013.

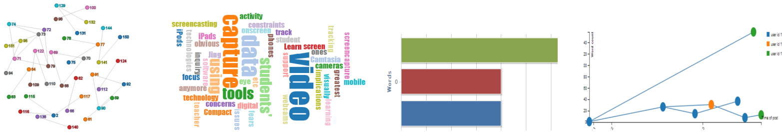


Fig. 1. Discussion visualisations: network graph, wordcloud, word count, thread plot

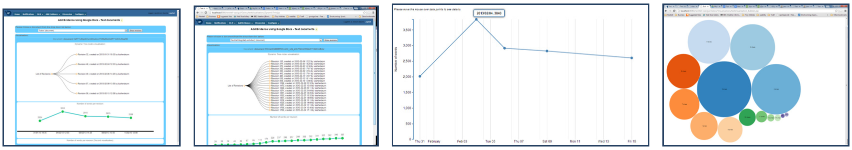


Fig. 2. Google Docs visualisations (revisions, semantic analyses between revisions; frequency)

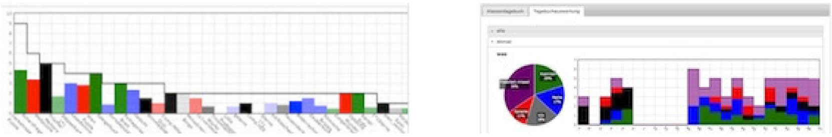


Fig. 3. Visualisations for materials used and time spent using them

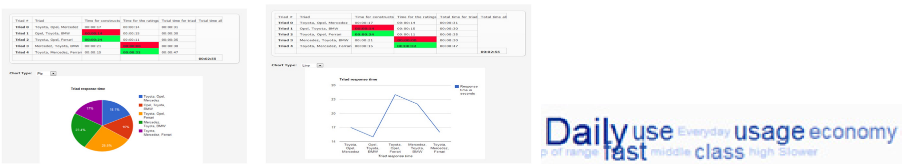


Fig. 4. Visualisations from a repertory grid exercise: response time; content word cloud

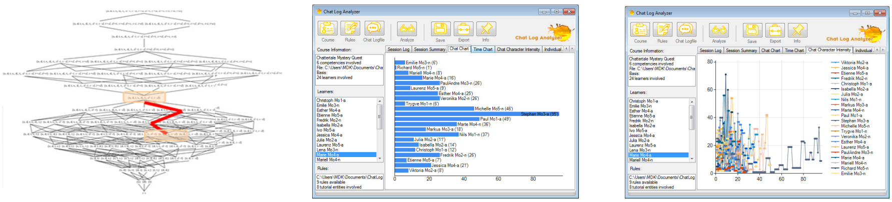


Fig. 5. Transforming detailed conceptual data for user interpretation

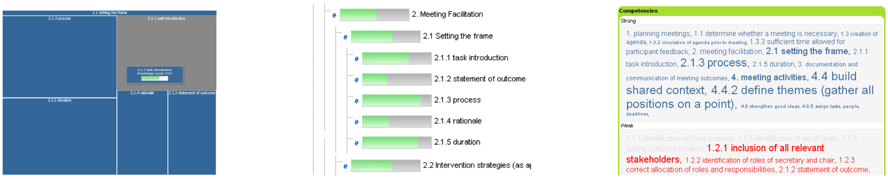


Fig. 6. OLM views: treemap; skill meters; word cloud

splitting of threads). Their benefits are correspondingly similar: users can see "what is going on". The same visualisations can be reused, allowing flexibility across activities, e.g. a teacher may wish to view data about a subject across different activity types, such as discussion forums, Google Docs, etc. Figure 2 gives additional examples for writing in Google Docs: revisions (including semantic analyses between revisions), and word frequency. As in Figure 1, these can also offer useful human-interpretable information about actions and progress. Figure 3 offers a different type of visualisation: learning materials used (left), with the dark line showing number of interactions with each material; and coloured blocks, the duration of interaction. The closer the block to the line, the closer the student is to completing the time for that topic (in a specific tool and context). Time spent is shown in the pie chart and calendar (right), to allow teachers to see whether students are interacting as expected.

Other approaches are more conceptually oriented. The repertory grid asks students to distinguish similar and different items from a group of three words/concepts. This generates activity data, but can additionally help teachers better understand their students, e.g. by identifying misconceptions if these are included in the exercise in the word groups presented. Figure 4 shows response time and a conceptual word could.

Each of the above, in or outside Next-TELL, can be effective. However, with increasing use of technologies in learning and increasing use of learning analytics come questions of how to use the visualisation tools. With more visualisation it may become less easy to gauge *learning*. In line with enhanced teacher understanding of students' knowledge or conceptualisations are approaches that model knowledge on a detailed level. The Next-TELL ProNIFA tool is based on Competence-based Knowledge Space Theory [4], which enables structuring and representation of a domain based on prerequisite relations using a set-theoretic framework (relations amongst problems such as test items). This can result in detailed domain representations and routes that an individual might take through a domain (see Hasse diagram left of Figure 5). This can be translated for teachers, and a variety of learning analytics produced (centre and right of Figure 5: from activity and chat logs in a virtual world). This interpretation requires pre-specified and defined conditions, heuristics and rules, but is a powerful approach when such specifications are available, or if a teacher wishes to define rules. In the case of the virtual world, it is difficult to get the data straight from the activity into a form where competencies can be easily recognised. Through its rules, ProNIFA allows activity and performance data to be displayed (Figure 5), but can also send it to the OLM. The OLM takes data from various sources (e.g. specific tools, discussion/chat interactions, Google Docs activity) and manual self, peer and teacher assessments. The data may come to the OLM through the API, if suitably structured, or it may be transformed using ProNIFA. The OLM is based on competency frameworks: teachers can select from existing frameworks (e.g. Common European Framework of Reference for Languages [5]), or can build their own, linking competencies with activities (e.g. competencies for meeting planning and facilitation in different activities: virtual world, chat, skype, face-to-face). Such a framework may span several subjects, e.g. 21st Century Skills, English, Business.

Figure 6 shows the treemap OLM view (of one of the sub-groups of a meeting competency framework); and corresponding information in skill meters and word cloud. This can be displayed to: the student that is modelled; peers if the model has been released; and teachers with reference to individuals or the group. Teachers may use this information in classroom orchestration if data comes from several sources, or for subsequent planning. Students may use the information primarily for metacognitive behaviours such as planning, self-monitoring, self-assessment, reflection – helping them take responsibility for some of the decisions in their learning. Users may drill-down to reach evidence for the learner model from a competency or an activity perspective. (Evidence may be quantitative; artefacts (e.g. essay, screen shot of learning analytics data); may comprise self or peer assessments; or automated data through ProNIFA or directly from another online activity.) By default, each activity contributing to the OLM is weighted equally, but more recent data has higher weighting. Teachers may alter these weightings. The OLM can therefore draw together data from various sources when the need is to focus on learner competencies, while other learning analytics visualisations are used to the extent that they best suit a specific purpose.

In summary: The challenge is when visualisations in different tools show different information in similar forms or, conversely, the same information in different forms. If teachers wish to benefit from the range of learning tools and activity tracking possibilities available, it is inevitable that they will encounter this. Even our own visualisations illustrate this (e.g. bar graphs for activities look similar to the OLM skill meters; word clouds show concepts in a repertory grid exercise and competencies in the OLM; pie charts show speed of response and materials used). Next-TELL does not expect all tools or visualisations to be used by a single teacher. Nevertheless, there may be several tools selected as applicable, and we expect that a teacher may also use other tools in addition to those offered by Next-TELL. Apparent 'clashes' will therefore likely occur. Given this, we suggest teachers use tools and learning analytics most suitable for the purpose at the time (essay revision visualisation when considering writing strategies; network graph for peer discussion; use of materials to indicate progress through tasks). The OLM can also take this data, and can be used on its own for recognising competencies with reference to individual activities as well as across activities. In some cases these visualisations may be used reflectively; in other cases they may be used to support on-the-spot decision-making (by learner or teacher).

Acknowledgement. This project is supported by the European Commission (EC) under the Information Society Technology priority FP7 for R&D, contract 258114 NEXT-TELL. This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made.

References

1. Long, P., Siemens, G.: Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE Review, 31–40 (September/October 2011)
2. Verbert, K., Duval, E., Klerkx, J., Govaerts, S., Santos, J.L.: Learning Analytics Dashboard Appliclations. American Behavioral Scientist (2013), doi: 10.1177/0002764213479363
3. Bull, S., Kay, J.: Open Learner Models as Drivers for Metacognitive Processes. In: Azevedo, R., Aleven, V. (eds.) Int. Handbook on Metacognition and Learning Technologies (2013)
4. Doignin, J., Falmagne, J.: Knowledge Spaces. Springer, Berlin (1999)
5. Council of Europe (not dated). The Common European Framework of Reference for Languages, `http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp` (accessed March 18, 2013)

ARLearn
Open Source Mobile Application Platform for Learning

Dirk Börner, Stefaan Ternier, Roland Klemke, Birgit Schmitz, Marco Kalz, Bernardo Tabuenca, and Marcus Specht

Centre for Learning Sciences and Technologies, Open Universiteit, Heerlen, The Netherlands
{dirk.boerner,stefaan.ternier,roland.klemke,birgit.schmitz, marco.kalz,bernardo.tabuenca,marcus.specht}@ou.nl

Abstract. The paper presents and outlines the demonstration of an open source mobile application platform for designing, supporting, and evaluating mobile learning scenarios that make use of media artefacts in a specific context. The platform contains a web-based authoring environment, cross-platform mobile applications to run the scenarios, as well as tools to monitor progress and results. Besides exploring the pedagogical background, the paper describes the conceptual implementation as well as the technical infrastructure and lists the requirements for demonstrating the platform and all its components.

Keywords: Mobile Learning, Serious Games, Open Source, Field Trip, Demonstration.

1 Introduction

As a result of on going technological developments and work on mobile learning applications, the authors' present ARLearn[1] - an open source mobile application platform for designing, supporting, and evaluating mobile learning scenarios that make use of media artefacts in a specific context. A flexible dependency mechanism enables the definition of an instructional (game) logic on top of these artefacts. The platform contains a web-based authoring environment, cross-platform mobile applications to run the scenarios, as well as tools to monitor progress and results. ARLearn offers great potential for different learning applications and has been successfully applied to support especially field trips and role-playing serious games.

2 Background

ARLearn allows defining instructional designs for mobile applications linking the virtual world with real world experiences. This concept is backed up by several educational theories. The anchored instruction approach [1] was developed to decrease the problem of inert knowledge through the presentation of real authentic problems and the active exploration by learners. ARLearn intends to implement this core idea in

[1] ARLearn: `http://code.google.com/p/arlearn/`

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 536–539, 2013.

linking real world situations and problems with learning support. Furthermore the theory of situated learning [2] is grounded on the assumption that learners do not learn via the plain acquisition of knowledge but they learn via the active participation in frameworks and social contexts with a specific social engagement structure. ARLearn facilitates such an authentic learning context by providing the means for immersive learning scenarios. According to [3] immersive learning is defined as learning that involves the "subjective impression that one is participating in a comprehensive, realistic experience". This covers potentially all phases of the experiential learning cycle [4], namely concrete experience, reflection, abstract conceptualization, and active experimentation.

2.1 Technology

The core of ARLearn is the capability to make use of media artefacts in a specific context. The platform allows defining a (game) logic on top of these artefacts. These artefacts can hold information or add a function to the game and can be positioned on a map by providing latitude and longitude attributes. Not providing location attributes turns the item into messages that users can receive at some point in time. Within a game, an author defines these items, as well as the following dependencies between the items:

- Action-based dependencies become true once a certain game action has been triggered,
- Time-based dependencies binds a time offset to another dependency, and
- Boolean dependencies provide means to create "AND" and "OR" statements with other dependencies.

Once a game has been created, an arbitrary amount of runs can be created and played. A run defines users grouped in teams. While users play a run, they generate actions (e.g. "reading a message", "answering a question") and responses. This output is then managed within the realm of a run.

The ARLearn platform is based on an application infrastructure. ARLearn builds on the JAVA version of Google App Engine[2] using servlet, Java Data Objects (JDO), JCache and other JAVA technology. Two communication protocols for clients permit retrieving information from ARLearn. The REST based Application Programming Interface (API) features basic CRUD (Create, Read, Update and Delete) operations. Furthermore a push notification system was implemented that only sends messages to the client when an event occurred. In this notification system, each client maintains a connection and listens for messages coming from the server. The platform is complemented by a web-based authoring environment, a cross-platform mobile applications to run scenarios, as well as tools to monitor progress and results.

[2] Google App Engine: `https://appengine.google.com`

3 Scope and Results

ARLearn has been successfully applied to support especially field trips and role-playing serious games. Respective results have been reported and published in workshop and conference proceedings as well as journals, e.g. [5] and [6]. Recently the platform is also used in the context of European projects on language learning with young children and to support inquiry-based learning of adolescents.

In total two papers related to the ARLearn application platform haven been submitted to the conference. The first submission describes the design of a mobile learning game to investigate the impact of role-playing on helping behaviour. The second submission used the platform to design a pervasive intervention to increase pro-environmental awareness, consciousness, and learning at the workplace.

4 Demonstration

In the context of the conference demonstration interested users will be able to explore all tools and existing scenarios created with the application platform, e.g. a field trip to explore cultural and architectural highlights of a city, a serious game for young children learning languages, a thematic tour in a museum, and a role playing scenario to train emergency situations. Besides the existing scenarios, a showcase explaining the functionality and approach of the platform itself as well as the related case studies and designs submitted to the conference can be experienced under guidance of the corresponding authors. Furthermore it will be possible to trial recent developments, such as mobile authoring or ambient information visualization.

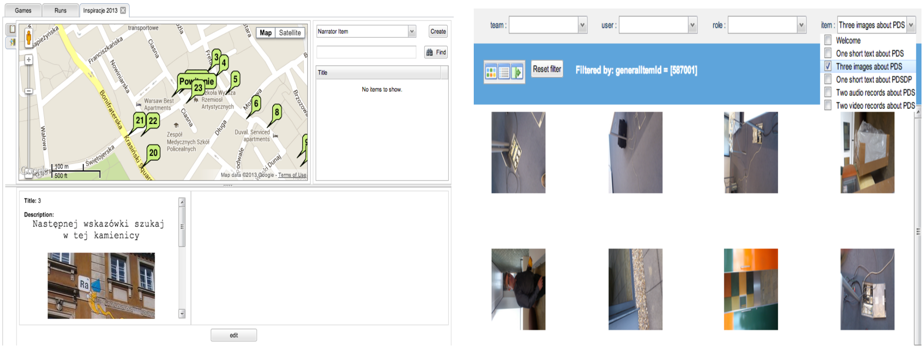


Fig. 1. ARLearn game authoring (left) and results display (right)

The demonstration will illustrate how authors can create new games (see Fig. 2 left). Via a web based authoring tool, users can create games and can add users to a game play. As soon as game players upload answers or data (e.g. pictures, video, etc.), one can use the ARLearn Results Display (see Fig. 2 right) to monitor student progress. This web based tool enables filtering as well as navigating through the results of a game play.

Fig. 2. ARLearn mobile application

The demonstration will showcase the different tools accompanied by a number of mobile devices to showcase the mobile application. The mobile application is currently only available for Android (Version 2.2 or higher). The application can be downloaded for free via the Google Play Store. In the context of the conference demonstration interested users can either download and install the application on their own devices or interact with the application on a limited number of demonstration devices. Figure 2 shows the mobile application for Android.

References

1. Bransford, J.D., Sherwood, R.D., Hasselbring, T.S., Kinzer, C.K., Williams, S.M.: Anchored Instructions: Why we need it and how technology can help. In: Nix, D., Spiro, R. (eds.) Cognition, Education and Multimedia: Exploring Ideas in High Technology, pp. 163–205. Erlbaum, Hillsdale (1990)
2. Lave, J., Wenger, E.: Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge (1990)
3. Dede, C.: Immersive interfaces for engagement and learning. Science 323(5910) (2009)
4. Kolb, D.A.: Experiential Learning: Experience as the Source of Learning and Development. Prentice-Hall, Englewood Cliffs (1984)
5. Gonsalves, A., Ternier, S., De Vries, F., Specht, M.: Serious games at the UNHCR with ARLearn, a toolkit for mobile and virtual reality applications. In: Specht, M., Sharples, M., Multisilta, J. (eds.) Proceedings of 11th World Conference on Mobile and Contextual Learning (mLearn 2012). CEUR Workshop Proceedings, pp. 244–247 (2012)
6. Ternier, S., Klemke, R., Kalz, M., Van Ulzen, P., Specht, M.: ARLearn: Augmented reality meets augmented virtuality. Special issue Technology for learning across physical and virtual spaces. Journal of Universal Computer Science 18(15), 2143–2164 (2012)

SynC-LD: Synchronous Collaborative IMS Learning Design Authoring on the Web

Michael Derntl, Petru Nicolaescu, Bezunesh Terkik, and Ralf Klamma

RWTH Aachen University
Advanced Community Information Systems (ACIS)
Informatik 5, Ahornstr. 55, 52056 Aachen, Germany
{derntl,nicolaescu,bezunesh,klamma}@dbis.rwth-aachen.de

Abstract. The IMS Learning Design (LD) specification enables the formal definition of teaching and learning flows. Several IMS LD authoring tools have been developed, most of them desktop based. There are few authoring tools that are deployed in a browser based environment, and some have built-in support for asynchronous collaboration during authoring via shared repositories. However, there are currently no tools available that enable synchronous, collaborative authoring in real-time. This demonstration presents SynC-LD, a novel widget-based tool that closes this gap. It supports browser-based, collaborative visual modeling of activity flows and the definition of IMS LD elements and their attributes. Multiple users can collaborate synchronously on the same learning design, which is achieved through inter-widget communication technology. Initial end-user trials show that the SynC-LD tool is usable and that IMS LD authors see potential in real-time IMS LD authoring.

Keywords: IMS Learning Design, real-time collaboration, widget-based user interface.

1 Introduction

IMS Learning Design (LD) is an interoperability specification [1] that allows learning designers to formally define activities, roles and resources in teaching and learning settings. Units of learning created and packaged this way can be exchanged between IMS LD authoring tools and deployed to compliant runtime tools. Since IMS LD was released in 2003 several authoring tools have been developed, for example Reload or ReCourse. Most of these tools are Java-based desktop applications. There are a few authoring tools available that can be used in a web browser, e.g., WebCollage or CADMOS. IMS LD authoring and available tools are discussed in more detail in [2].

Some of the available tools support sharing, publishing, retrieval, and repurposing of the created learning designs. Combining these steps allows asynchronous collaboration on these units of learning, similar to file sharing in a shared file space. However, there are currently no tools available that enable

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 540–543, 2013.

synchronous collaborative authoring in real-time. The SynC-LD application presented in this demo enables just that: real-time authoring of IMS LD units of learning in a web browser by two or more authors simultaneously.

2 Technical Approach

The SynC-LD application is offered as a web widget—that is, a reusable component running in a web browser and offering specialized functionality. Widgets support a wide range of devices and multiple widgets can be mashed up into portal containers to create personalized computing and work environments. The SynC-LD widget can be used together with other widgets like chat, sketching tools, etc., to improve the collaboration experience.

As an environment for developing and deploying SynC-LD, the ROLE SDK[1] was chosen, since it offers the basic infrastructure components and settings that are necessary for collaboration, messaging, and inter-widget communication. In a ROLE space, all widgets are rendered in the browser using an Apache Shindig container[2]. SynC-LD runs in any ROLE space, which is an abstract concept of a Personal Learning Environment modeling groups of users and enabling interaction among them. Using the Extensible Messaging and Presence Protocol (XMPP), real-time communication between users (e.g., membership, multi-user chat) or between widgets (inter-widget communication, enabling the exchange of information between widgets across browsers) can be achieved. The ROLE inter-widget communication (IWC) framework is an implementation of the XMPP Publish-Subscribe extension protocol, which is also frequently used in developing multi-user chat applications.

The SynC-LD widget uses the widely implemented web languages Javascript, HTML5 and CSS. In order to ensure synchronization of operations (editing of attributes, adding and removing elements) across all clients, the Operational Transformation (OT) [3] technique is used. OT supports real-time collaborative editing for multiple users, without imposing restrictions such as locking, where only one client can edit at a given moment. It maintains the document's consistency and ensures resolution of conflicting operations and intention preservation. To achieve these requirements, each participating site has its own instance of the OT engine. The local changes are immediately applied and sent to the remote clients unchanged, whereas the remote operations are first transformed before they are applied. Therefore, each operation is first checked for conflicts before being reflected into the IMS LD structure. Furthermore, from the architectural design perspective, the prototype is realized for peer-to-peer collaboration. As such, SynC-LD does not require any central server. The communication between clients (operations, synchronization messages, etc.) is realized using the ROLE IWC, in a simple format, similar to Google Android intents.

Since IMS LD is an XML-based specification, we use an object identifier generator to distinguish between the different XML elements which need to

[1] `https://github.com/ROLE/ROLE/wiki/ROLE%20SDK`

[2] `http://shindig.apache.org/`

be synchronized. The OT library ensures that the text values of attributes, check boxes and the different objects added or removed from the description are synchronized across all peers, i.e., all the clients have the same view on the data.

3 Results

The resulting web application allows users to create all required elements that make up a valid IMS LD unit of learning. The metaphor for visual editing of activity flows is similar to the metaphor used in the OpenGLM authoring tool [4]. In Fig. 1, for example, we see a model arranging four learning activities in a sequence, which is achieved by dragging and dropping the symbols in the "Tools" bar onto the modeling canvas and connecting them with transition arrows.

The other tabs in the SynC-LD user interface allow specifying the required elements defined in level A of IMS LD. This includes under the "Properties" tab the complete details of the drawn activities. A cut-out of the description of the "Read material" activity is shown in Fig. 2. The remaining tabs allow users to define "Environments" referenced in the activities, including learning objects and services; the "Roles" involved in the unit of learning (learners and staff); the "Role-parts" connecting the roles with activities; and various types of "Resources" used in other elements of the design (e.g., web sites or local files).

Different authors can simultaneously edit the same or different portions of the design. Every single change (even single characters typed into text boxes, or checks made in checkboxes) is immediately synchronized and visualized in the interface of all users participating in the authoring session. Finally, by clicking the "Export Design" button any user can export a zipped IMS LD package to the local machine at any time.

End-user evaluations of SynC-LD are currently in progress. Four guided authoring sessions with four distributed, collaborating users in each session so far have produced positive feedback in terms of usability and usefulness of the tool. All these users agreed or strongly agreed that synchronous collaboration on learning designs is a useful feature. Improvement suggestions were mostly

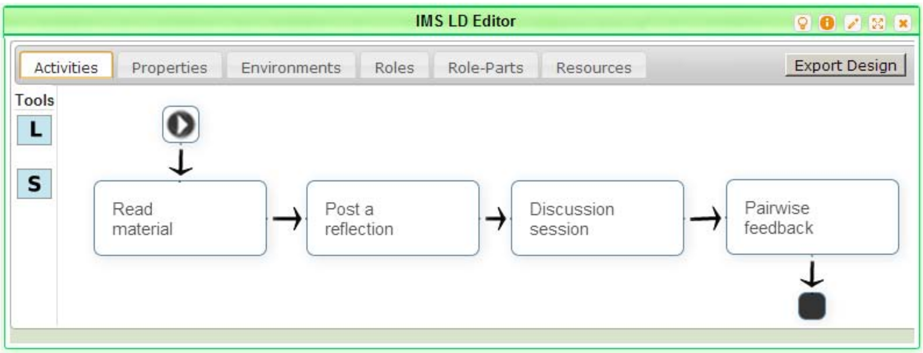


Fig. 1. Modeling canvas

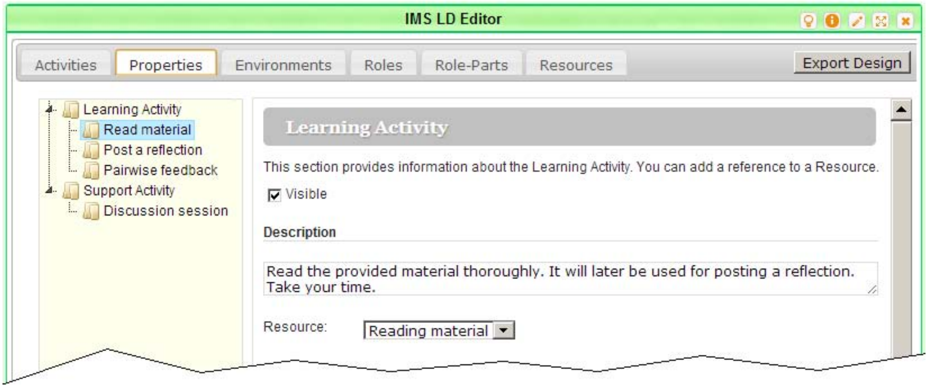


Fig. 2. Activity property editor (cut-out)

related to improved awareness support to see what portions of the design others are currently working on.

4 Demonstration

The demonstration will include the presentation of the running prototype through a brief learning design activity. The presenter will initiate a collaborative course design session and with the help of (invisible) collaborators the presenter will be able to magically complete the design while presenting the tool features and ideas to the audience. Changes by the collaborators will be visible instantly on the main projector screen. This way the key feature of SynC-LD, i.e., the real-time learning design authoring, can be experienced live while getting an explanation of how it works under the hood.

Acknowledgments. Development of SynC-LD and its underlying technologies was supported by the European Commission in the FP7 integrated projects LAYERS (318209) and ROLE (231396), as well as in the Lifelong Learning Programme multilateral project METIS (531262-LLP-2012-ES-KA3-KA3MP).

References

1. IMS Global: IMS Learning Design Specification (2003), `http://is.gd/imsldim`
2. Martínez-Ortiz, I., Sierra, J.L., Fernández-Manjón, B.: Authoring and Reengineering of IMS Learning Design Units of Learning. IEEE Transactions on Learning Technologies 2(3), 189–202 (2009)
3. Ellis, C.A., Gibbs, S.J.: Concurrency control in groupware systems. In: Proc. of ACM SIGMOD 1989, pp. 399–407. ACM, New York (1989)
4. Derntl, M., Neumann, S., Oberhuemer, P.: Propelling Standards-based Sharing and Reuse in Instructional Modeling Communities: The Open Graphical Learning Modeler (OpenGLM). In: Proc. of ICALT 2011, pp. 431–435. IEEE (2011)

Blending Evidence and Users for TEL: An Overture

Tania di Mascio[1], Rosella Gennari[2], and Pierpaolo Vittorini[3]

[1] DIEI – University of L'Aquila – Via G. Gronchi – 67100, L'Aquila, IT
[2] CS Faculty, Free University of Bozen-Bolzano – P.za Domenicani, 3 – 39100 Bolzano, IT
[3] MISP – University of L'Aquila – P.le S. Tommasi, 1 – 67100 Coppito, L'Aquila, IT

Abstract. TERENCE is an adaptive learning system for reasoning about stories with children having deep text comprehension problems. It develops reading interventions in the form of smart games for stimulating the text comprehension of such children. In order to ensure the pedagogical effectiveness and the usability of the smart games, and of the system in general, TERENCE was designed combining the user centred and the evidence based design. In this paper, we illustrate how such methodologies were used in the design of the TERENCE smart games.

1 Introduction

Developing the capabilities of children to comprehend written texts is key to their development as young adults. From the age of 7–8 until the age of 11, children develop as independent readers. Nowadays, more and more children in that age range turn out to be poor (text) comprehenders: they demonstrate difficulties in deep text comprehension, despite well developed low-level cognitive skills like vocabulary knowledge.

TERENCE is a Collaborative Project funded by the EC under the ICT Call 5 FP7-ICT-2009-5 (1 October 2010-30 September 2013) and aims at designing and developing an intelligent *adaptive learning system* (ALS) with learning material adequate for poor comprehenders [3]. Therefore the types of users of the TERENCE ALS are: *learners*, namely, primary-school poor comprehenders, hearing and deaf, older than 7–8; *educators*, namely, primary-school teachers and support teachers, as well as parents of the TERENCE learners. The learning material of the TERENCE ALS is in English and in Italian. The material is made of stories, collected into books and divided into difficulty levels for the TERENCE learners, and of smart games for each story, centred around inference-making interventions necessary for text comprehension. The main learning tasks of the learners with TERENCE are reading the adequate level of stories and playing with the adequate level of smart games for each story. Reading and playing are organised in cycles according to the stimulation pedagogy plan of TERENCE [5].

The TERENCE system and its learning material are developed following the user-centred design (UCD) [8] and the evidence-based design (EBD) [2]. Generally speaking, the UCD places the user, actively, at the centre of the design process, which is iteratively repeated until attaining the usability of the system. The EBD of a system stresses the need of basing design decisions on empirical evidence, involving experts of the domain in the process. In TERENCE, the UCD and EBD were mixed for analysing the context of use of the system and evaluating incremental versions of the systems.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 544–547, 2013.

This paper illustrates how such methodologies were used for iteratively and incrementally developing the TERENCE smart games, starting from the context of use analysis and ending with their evaluations.

The paper sketches the data gathering for setting the requirements for the smart games in Sect. 2 and the resulting design of the TERENCE smart games in Sect. 3. Sect. 4 presents the evaluations with experts or users, and their aims. This paper ends with a recap conclusive section.

2 The Context of Use Analysis for Smart Games

The context of use of TERENCE was analysed for specifying the characteristics of the users, the learning tasks the system should support and environment constraints. In the remainder, we focus on the learning tasks.

The main data for specifying the learning tasks and hence designing the smart games of TERENCE were gathered through

- contextual inquiries with experts of (poor) text comprehension,
- contextual inquiries with educators,
- and game-like field studies with learners.

See [9]. In particular, the data gathered from experts are, mainly, relevant reading interventions and how they should be scheduled for effectively stimulating the text comprehension of the TERENCE learners. For instance, according to the experts responsible for the pedagogy plan of TERENCE, rewards are a valuable mechanism for motivating the learners to continue playing and allow for breaks; relaxing games, stimulating visuo-perceptual skills and not reading comprehension, should have precisely such goals and work as rewarding mechanism after playing with the more demanding smart games. On the other hand, said the domain experts, smart games should be designed like interventions for the analysis of stories [9], supported by adequate visual representations, so as to guide the child to better recall and correlate the information that are acquired reading a story. More precisely, smart games should stimulate reasoning about: the characters and and their participation in the stories; temporal relations between events; causal-temporal relations between events. The interventions selected for the smart games were thus hierarchically organised and scheduled into macro-levels according to their main pedagogical goal:

1. at the entry macro-level, interventions focusing on characters, layered into two levels: who the agent of a story event is (who), what a character does in the story (what);
2. at the second macro-level, interventions focusing on time, layered into four levels: for reasoning about temporal relations between events of the story, purely sequential (before-after) or not (all the others);
3. at the last macro-level, interventions focusing on causality, layered into three levels: the cause of a given event (cause), the effect (effect), or the cause-effect relations between two events (cause-effect).

The TERENCE smart games were then layered into similar levels. The following section sketches how the design of smart games was conducted and based on the requirements established by the domain experts.

3 The TERENCE Smart Game Design

According to [1], a game should specify the following data: the *instructions* and overall *goal* of the game, the *initial state* of the game, the *termination* state, the legal *actions* of the players, and the *maximal duration time* per action if foreseen. For specifying the data for the TERENCE smart games, we analysed the requirements illustrated above. Then we abstracted and structured the data into the TERENCE game framework. The TERENCE framework was then specialised for each game level, e.g., we have a specific framework for all before-after game instances.

The TERENCE game framework is based on similar frameworks found in the literature, e.g., see [6]. However, the TERENCE framework is for puzzle or casual games like in TERENCE, and hence it is more specific and structured, lending itself to the automated generation of textual smart games by mixing constraint-based reasoning and natural language processing technologies, see [4] for an outline of the overall generation process, and [7] for the resulting textual instances of smart games.

The TERENCE framework of each game level was then used to create a related visual template. This was used for designing the interface of each game level, according to the age of the TERENCE learners. The interface design and its evaluation is briefly sketched below.

4 The Evaluations of Smart Games for Usability and Pedagogical Effectiveness

The TERENCE smart games were released in three main versions: the first was released in March 2012; the second was released in June 2012; the third was released in March 2013.

The first version implemented 9 smart game instances, at least one per smart game macro-level—character, time, causality. The prototypes of smart games were evaluated via heuristic evaluations or expert evaluations by experts of text comprehension or of interaction design. One of the main goals of the evaluations was to assess whether the smart games could be pedagogically effective for the TERENCE learners according to experts of text comprehension that were not involved in the context of use analysis. The other main goal was to detect potential usability issues with the graphical user interface with experts of interaction design.

The second version had games for two stories, with circa 12 smart games per story. The games were evaluated with 168 TERENCE learners in order to get pedagogical indications concerning the effectiveness of the TERENCE smart games and detect usability issues for the learners. The data collected were quantitative, e.g., performance metrics on tasks, and qualitative, mainly gathered via observations or interviews.

The third version of the TERENCE system implemented all the TERENCE stories, which are 32 written into 4 levels of difficulty and organised into books. For each story, this version of the system has circa 12 smart games, for a total of more than 380 smart games. This version is currently used in a long-scale evaluation for completing the study of the pedagogical effectiveness of the system and of its usability. The evaluation involves more than 900 learners across Italy and UK; it uses standardised tests of text comprehension for assessing the pedagogical effectiveness of the system as well as standard measures of usability, like time of tasks.

5 Conclusions

TERENCE is an European project developing an ALS specific for poor comprehenders and their educators, by combining the UCD and EBD methodologies. The project, thus, seeks the constant involvement of the users (poor comprehenders, deaf children and their educators) from schools in the UK and in Italy and domain experts, in particular, of text comprehension. Having both the users and domain experts involved in the context of use analysis and evaluation stages of the project, iteratively, helps to early detect usability and pedagogical effectiveness issues and solve them in time. In this paper, we outlined how the smart games of TERENCE were designed following the UCD and EBD methodologies, starting with the context of use analysis, and stepping through the design of the TERENCE smart games that was iteratively and incrementally refined via evaluations. To the best of our knowledge, TERENCE is the first technology enhanced learning project that combined UCD and EBD in such a manner for attaining the usability and pedagogical effectiveness of its system.

Acknowledgments. The authors' work was supported by TERENCE project. TERENCE is funded by the European Commission through the Seventh Framework Programme for RTD, Strategic Objective ICT-2009.4.2, ICT, Technology-enhanced learning. The contents of the paper reflects only the authors' view and the European Commission is not liable for it. The second author's work was also supported by the DARE project, financed by the Province of Bozen-Bolzano.

References

1. Adams, E.: Fundamentals of Game Design. New Riders (2010)
2. Becker, F., Hamilton, K., Zimring, C., Shepley, M.: A Practitioner's Guide to Evidence-Based Design. The Center for Health Design (2008)
3. Brusilovsky, P.: Adaptive Hypermedia. User Modeling and User-Adapted Interaction 11, 87–110 (2001)
4. Cofini, V., De la Prieta, F., Di Mascio, T., Gennari, R., Vittorini, P.: Design Smart Games with Context, Generate them with a Click, and Revise them with a GUI. Advances in Distributed Computing and Artificial Intelligence Journal (December 3, 2012)
5. Cofini, V., Di Giacomo, D., Di Mascio, T., Necozione, S., Vittorini, P.: Evaluation Plan of TERENCE: When the User-Centred Design Meets the Evidence-Based Approach. In: Vittorini, P., Gennari, R., Marenzi, I., de la Prieta, F., Rodríguez, J.M.C. (eds.) International Workshop on Evidence-Based TEL. AISC, vol. 152, pp. 11–18. Springer, Heidelberg (2012)
6. Consortium EMAPPS. EMAPPS Game Framework, `http://emapps.info/eng/Games-Toolkit/Teachers-Toolkit/Games-Creation/Framework-for-Game-Design` (retrieved January 2012)
7. Gennari, R.: Generation Service and Repository of Textual Smart Games. Technical report, TERENCE project (2012)
8. Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., Cajanger, A.: Key Principles for User-centred Systems Design 22(6), 397–409 (2003)
9. Slegers, K., Gennari, R.: Deliverable 1.1: State of the Art of Methods for the User Analysis and Description of Context of Use. Technical Report D1.1, TERENCE project (2011)

An Intelligent Tutoring System to Evaluate and Advise on Lexical Richness in Students Writings

Jesús Miguel García Gorrostieta[1], Samuel González López[2]
Aurelio López-López[2], and Maya Carrillo[3,⋆]

[1] Universidad de la Sierra, Moctezuma, Sonora, México
jmgarcia@unisierra.edu.mx
[2] Instituto Nacional de Astrofísica, Óptica y Electrónica, México
{sgonzalez,allopez}@inaoep.mx
[3] Facultad de Ciencias de la Computación, BUAP, México
cmaya@cs.buap.mx

Abstract. Lexical competence, the writer ability to use properly vocabulary, becomes a basic issue of a writing instructor when reviewing drafts. Here, we present the basic part of a web-based intelligent tutoring system to provide student guidance and evaluation in structuring research proposals. We elaborate a network-based model to follow the progress of each student in the development of the project, supply assignments and personalized feedback on each evaluation. This tutor includes for now a module for assessing the lexical richness, in terms of three measures: variety, density, and sophistication, that are described. We also explain the methodology for pilot testing with undergraduate students, whose results were encouraging, indicating that the tutor indeed helps students.

Keywords: intelligent tutoring system, lexical richness, density, variety, sophistication.

1 Introduction

The guidance of students on writing is a hard and time consuming task for advisors, since requires several iterations to achieve an acceptable level. There is a need to alleviate such burden, possibly by the use of tutoring systems.

An intelligent tutoring system (ITS) offers personalized instruction or feedback to students without much involvement of instructors. Recent advances include natural language technics to analyse essays, interact, and provide feedback, such as Writing Pal [1], intelligent virtual agents [2], Guru [3], or Auto Tutor [4].

Lexical competence, i.e. the writer ability to use properly vocabulary, is an essential factor affecting writing, being considered a basic reference point for measuring the quality of writing [5]. Thus, any assessment of writings has to start from this competence. We view lexical richness as a way to hint on it.

⋆ Second author was supported by Conacyt (scholarship 1124002), while third and fourth authors were partially supported by SNI, México.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 548–551, 2013.

A study [6], aimed to relate the academic success of students with lexical richness, used as parameters vocabulary size and lexical variety, among others. They conclude that the results suggest that students with appropriate vocabulary, varied and accurate, have excelled in their studies, while students with a general repetitive and uncontrolled vocabulary showed a decreased performance.

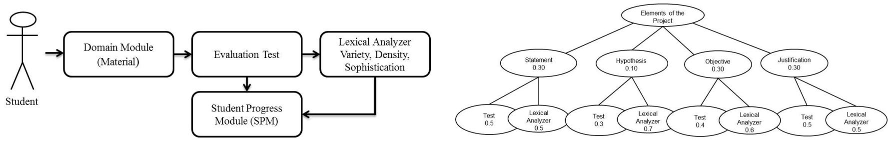
There are several methods to evaluate the use of vocabulary. One is to measure sophistication of text using word lists, e.g. [7] used a list of 3000 easy words. For Spanish, some studies use the list provided by the Spanish Royal Academy (SRA) of most frequent words.

When advising students, the correct wording or abuse of certain terms become an issue for the instructor. Furthermore, the process of drafting a proposal is usually not an easy task for students. Hence, our proposed system intends to assist the work of the instructor and to facilitate and guide students. Here, we present the basic part of a web-based ITS [8] to provide student guidance and evaluation in structuring research projects. We propose a network-based model to follow the progress of each student in the development of the project and personalized feedback on each assessment. This tutor includes for now a module for assessing lexical richness in terms of three measures: variety, density, and sophistication. We perfomed additionally pilot testing with undergrad students.

2 The Model

The intelligent tutor in the Domain Module (Fig. 1(a)) presents material concerning the different elements of the project, such as the problem statement, hypothesis, objectives and justification. A test is applied to validate the reading of materials of each element and then practical exercises are applied using the Lexical Analyzer to achieve a high level of richness measures. The results of the test and analysis are sent to the Student Progress Module (SPM) to update the student knowledge state. SPM records the student progress in a network representation (Fig. 1(b)). When the student completes the test, the value of the test node element is updated and SPM calculates the student progress for the parent node, considering the weights assigned to each question in the test. Likewise, when doing the exercises with the Lexical Analyzer, the corresponding node is updated and the SPM estimates the student progress for the parent node using the weights assigned to the measure in turn. Weights were assigned to each node based on instructors experience. Finally, the results of the Lexical Analyzer are sent to the SPM, to manage the student achieved results.

Lexical analysis focuses on the evaluation of three measures: variety, density, and sophistication, which together assess lexical richness. Variety seeks to measure student ability to write with a diverse vocabulary, and is computed by dividing the number of unique lexical types (terms of content) by the total of lexical types, both ignoring empty words. Density aims to reflect the proportion of content words respect to the complete text, intended to reveal an excessive use of empty words, and is calculated by dividing the number of unique lexical types or content words by the total words of the evaluated text (i.e. before removing stop words). Sophistication attempts to reveal the handling of technical



(a) General Model. (b) Network of Student Model.

Fig. 1. Intelligent Tutoring System

concepts and is the proportion of "sophisticated" words employed, and is computed as the percentage of words out of a list of 1000 common words, according to SRA. A scale ranging in High, Medium and Low in lexical measures has been set based on previous work [8], where we analyzed research proposals and theses of different levels.

3 Intelligent Tutoring System and Pilot Testing

The implemented ITS (Fig.2, in Spanish) has a menu on top that allows to access the elements of the project, i.e. problem statement, hypothesis, objectives, or justification. For each element, there are three sections: material, test, and practical evaluation. There is also the progress section in the left side, to report the progress. To enter the practical evaluation, the student must first complete well the test receiving a 50% of advance in the concept and 15% in the complete project. In the section of practical evaluation, the student enters a problem statement to be analyzed. Once completed the three lexical analyses, the student can move on to the next item of the project draft.

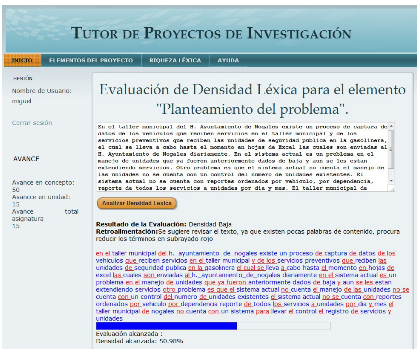


Fig. 2. ITS Interface (in Spanish)

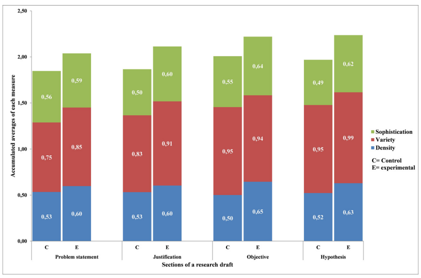


Fig. 3. Results of Pilot Test

We performed a pilot testing to assess the impact of using an ITS focused on lexical richness in elements of a research project. This experiment involved undergrad students of Telematics program, divided in two ramdomly selected groups;

(E)xperimental and (C)ontrol, both of 14 students in a Simulation course, where they had to write a research proposal. Figure 3 shows the average of the measures obtained by both groups. Group E produced higher averages than group C, for the different sections, reaching the highest value in *Hypothesis* with 2.24 and the lowest in *Problem Statement* with 2.04, both of group E. For group C, we found the values of 2.01 for *Objective* and 1.85 for *Problem Statement*. From a statistical hypothesis test, we can state that the ITS helps students to improve the three lexical aspects.

4 Conclusions

The use of an ITS for research project drafts aims to support teachers in reviewing research proposals providing material to the student, by tracking their progress and lexically analyzing their writings. The pilot test done with two groups of students allow to have some evidence that the use of ITS including the lexical analysis help students improve their writings in terms of their lexical richness as measured. We are in the process of including additional sections for analysis in the tutor, to support a more elaborated structure of drafts. Lexical analysis will be complemented with other analysis, such as the evaluation of coherence and language models by sections.

References

1. McNamara, D., Raine, R., Roscoe, R., Crossley, S., Jackson, G., Dai, J., Cai, Z., Renner, A., Brandon, R., Weston, J., Dempsey, K., Carney, D., Sullivan, S., Kim, L., Rus, V., Floyd, R., McCarthy, P., Graesser, A.: The Writing-Pal: Natural language algorithms to support intelligent tutoring on writing strategies. In: Applied Natural Language Processing and Content Analysis, pp. 298–311. IGI Global, Hershey (2012)
2. Rospide, C.G., Puente, C.: Virtual Agent Oriented to e-learning Processes. In: Procs. International Conference on Artificial Intelligence, Las Vegas, NV (2012)
3. Olney, A.M., D'Mello, S., Person, N., Cade, W., Hays, P., Williams, C., Lehman, B., Graesser, A.: Guru: A Computer Tutor That Models Expert Human Tutors. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 256–261. Springer, Heidelberg (2012)
4. Graesser, A., D'Mello, S., Craig, S., Witherspoon, A., Sullins, J., McDaniel, B., Gholson, B.: The Relationship between Affective States and Dialog Patterns during Interactions with Autotutor. J. Interactive Learning Research 19(2), 293–312 (2008)
5. Grobe, C.: Syntactic Maturity, Mechanics, and Vocabulary as Predictors of Quality Ratings. Research in the Teaching of English 15(1), 75–85 (1981)
6. Roy, S.: Non-Native English Speaking Students at University: Lexical Richness and Academic Success. Doctoral Thesis, University of Calgary, Canada (2010)
7. Schwarm, S., Ostendorf, M.: Reading level assessment using support vector machines and statistical language models. In: Procs of the 43rd Annual Meeting on Association for Computational Linguistics (ACL 2005), pp. 523–530 (2005)
8. González, S., López-López, A.: Supporting the Review of Student Proposal Drafts in Information Technologies. In: Procs. ACM SIGITE 2012 & RIIT, pp. 215–220 (2012)

A Tool to Aid Institutions Recognize Their Employees Competences Acquired by Informal Learning

Francisco J. García-Peñalvo, Valentina Zangrando, Alicia García-Holgado, Miguel A. Conde-González, Antonio-M. Seoane-Pardo, Marc Alier, Nikolas Galanis, Jordi López, José Janssen, Francis Brouns, Anton Finders, Adriana Berlanga, Peter Sloep, Dai Griffiths, Mark Johnson, Elwira Waszkiewicz, Aleksandra Mykowska, Miroslav Minovic, Milos Milovanovic, Maria A. Marques, Maria C. Viegas, and Gustavo R. Alves

The TRAILER Project
grial@usal.es
http://grial.usal.es/agora/trailerproject/

Abstract. People do not learn only in formal educational institutions, but also throughout their lives, from their experiences, conversations, observations of others, exploration of the Internet, meetings and conferences, and chance encounters etc. However this informal and non-formal learning can easily remain largely invisible, making it hard for peers and employers to recognize or act upon it. The TRAILER project aims to make this learning visible so that it can benefit both the individual and the organization. The proposed demonstration will show a software solution that (i) helps the learners to capture, organize and classify a wide range of 'informal' learning taking place in their lives, and (ii) assists the organization in recognizing this learning and use it to help managing human resources (benefiting both parts). This software tool has recently been used in two phases of pilot studies, which have run in four different European countries.

Keywords: non-formal learning, informal learning, tagging competences, software tool.

1 Background (Pedagogy)

The Information & Communication Technologies (ICT) advances cause knowledge management flows about innovation-acceptation-consolidation-obsolescence that regard to both institutions and individuals. In learning contexts, ICT has changed the perspective from the different educational stakeholders that demand an increase of personalization, more connectivity with other peers, unlimited access to resources and information sources, a plenty flexibility in the way, place and time they access, and a natural and necessary coexistence of the formal and informal learning flows. Nowadays, this convergence among formal, non-formal and informal learning causes an international debate about the validity and evolution of the traditional teaching and learning models, especially in regard to the professional training processes [1, 2].

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 552–555, 2013.

Informal learning takes place in the context of everyday experience in both workplace and higher education, emerging from the activity rather than being planned. Developing the digital literacy skills needed to be part of this participatory culture is a key challenge facing education today. These skills are way beyond simple notions of ICT literacies and are more about harnessing the affordances of social and participatory media.

One of the main challenges regarding informal learning is the recognition of the developed competences throughout these informal channels. This is especially relevant because it enhances and produces positive benefits for managers and companies; it may develop task skills and know-how and communicates "social" norms and preferred patterns of behavior; it gives employees the opportunity to learn and keep their skills up-to-date, while being part of the overall workplace culture rather than just its training regime, etc. [3]. In this context, TRAILER represents a dialog layer between the institution and the involved persons (informal learners) allowing informal competences identification, tagging, sharing and recognition for a symbiotic exploitation of this unknown knowledge of both institutional and personal development.

2 Background (Technology)

TRAILER defines a methodology to make informal learning experiences transparent to learners and institutions so that both sides may benefit. The methodology comprises a framework, described in Fig. 1, with several components and interfaces to make possible the interaction required. A personal learning environment groups the tools that the user interacts with in their informal learning [4]. These could include Social Networks, LMS, and Remote Labs, among others. One of the tools included is the TRAILER portfolio that has an interface called Informal Learning Collector (ILC) to facilitate the gathering of informal learning activities (ILA). In addition, there are several institutional tools like: a Competence Catalog that facilitates a way to categorize informal learning experiences taking into account learner or institutional perspectives; an Institutional Environment that facilitates the analysis of the published information in order support dialog with the learner and to facilitate decision-making concerning learning issues within the institution (e.g., accreditation processes); and a Repository to store the information to be analyzed and facilitate the generation of reports that could be useful to both institution and learners. Taking into account that the ILA can be carried out with very different tools (defined in different programming languages and with different data formats), it is necessary to design a service-based architecture to support this diversity. In such architecture the PLN includes browser-based activities, widgets, games, remote labs; the ILC is based on Moodle; the ePortfolio is developed using Liferay; and the institutional environment on a PHP development. These components exchange information through REST and JSON-RPC interfaces.

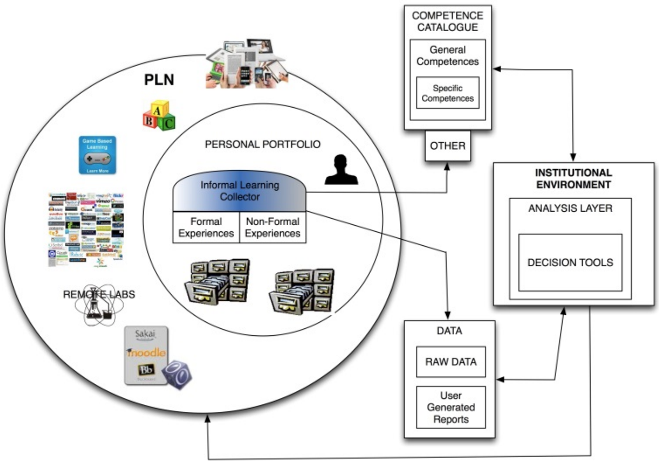


Fig. 1. Description of TRAILER components

3 Results and Outcomes Achieved

TRAILER project pilots were implemented in four countries, both in a few chosen companies and academic institutions (higher education).

Pilots have two distinct and asynchronous phases: Pilot 1 is focus on learners while Pilot 2 is focus on institutions. Pilot 1 began with a hand-on workshop (WK) where the TRAILER project was presented and the tool was explained step by step, each followed by proposed tasks. Participants were then asked to try the tool for one week, keeping in mind the performance of some particular tasks: collecting evidences, defining competences and organizing their information in showcases. Finally they were invited to participate in a Focus Group session (FG) where they could share their opinions and make suggestions. Before the WK and the FG session, participants had to answer an on-line questionnaire (pre & post), related to their informal learning (IL) perceptions.

After analyzing the reports on how the WK was received by the learners and also their reactions, problems and comments, it is possible to point out some common features and some particular aspects that might have conditioned their participation. The learners groups were somehow different namely concerning motivation in using this type of tool. Also similar doubts and concerns show common features that must be taken in consideration, namely: *Why the ILC and the Portfolio are separate? Why is not possible to delete an activity from the Portfolio? Exactly what is made public?*

4 Demonstration Outline

The demonstration will follow the two-phase process already described; starting by addressing the learners in capturing and organizing their IL, and finishing demonstrating how the organizations may recognize this learning and what they can benefit from. In a first stage, learners are asked to collect evidences of their IL activities and associate them with competences. These competences include those defined within a general list of competences (based on ISCO-88) and those previously defined by the organization as competences relevant to their own business activities. As the learners can collect IL evidences in a random manner, the Portfolio provides a way to organize these activities in showcases and decide which are to be shared with others (within interest groups: professional, personal, or others). Learners may always decide whether or not to publish a competence, which, by doing so, becomes visible to their superiors. Also, the latter might be interested in knowing their learners competences, in order to support decisions that benefit both. This information is available to them through their institutional environment. Based on the visible information, better sustained decisions, like *"Has the organization competences to endeavor a certain project?"*; *"Has the organization the need for external contracting for a specific task?"*; *"How the organization can manage the learners' activities to their best interest?"* are expected to be addressed using the TRAILER tool. This tool also provides a way to find people with similar competences within the TRAILER community, which represents an opportunity to exchange ideas and to explore collaboration opportunities.

References

1. Cross, J.: Informal learning: Rediscovering the natural pathways that inspire innovation and performance. Pfeiffer, San Francisco (2007)
2. Kamenetz, A.: Edupunks, edupreneurs, and the coming transformation of higher education. Chelsea Green Publ., White River Junction (2010)
3. Halliday-Wynes, S., Beddie, F.: Informal learning: At a glance. National Centre for Vocational Education Research, Adelaide (2009)
4. García-Peñalvo, F.J., Conde, M.Á., Alier, M., Casany, M.J.: Opening Learning Management Systems to Personal Learning Environments. Journal of Universal Computer Science 17, 1222–1240 (2011)

Engaging Learning Technologies for the Classroom of Tomorrow

Fabrizio Giorgini[1] and Peter Reimann[2]

[1] Lattanzio Learning, Via Portobello, 16039 Sestri Levante, Italy
giorgini@lattanziogroup.eu
[2] MTO, Tübingen/Germany
p.reimann@mto.de

Abstract. While the technology-rich classroom makes it comparatively easy to gather, store and access data on students' activities, turning those into information on learning that can inform pedagogical decision-making is still hard to achieve. In the NEXT-TELL project, we build on concepts from educational assessment design and on modeling concepts from computer science as a basis for generating quality data on students' learning. We describe a set of inter-related methods and software components that can be used to turn assessments into support mechanisms for learning, and to make use and sense of data for improving teaching and learning.

Keywords: technology-enhanced learning, future, formative assessment, learning analytics.

1 Introduction

Schools are slowly yet inevitably entering the information age. But while the level of technology infusion is increasing, and with it the capacity to distribute information for learning and gather information about learning quickly and efficiently, we are still far away from the vision of the school as a "high performance, personalized learning community" [1]. The barrier to that is not so much the absence of information, but the absence of *the right information, at the right time, in the right format*. The classroom may be increasingly data-rich, but is still comparatively information-poor. One reason is that a good part of the data is usually not linked to classroom practices and outcomes, and is not available close enough in time to learning and teaching activities [2].

An important step for making classroom data useful for teachers' and students' 'tactical' decision making ('what to do next?') is to express the information on the level of knowledge and skill development, processes of learning, motivation and engagement. Teachers usually get this information from direct observations, and from formative assessments, such as quizzes or problem solving exercises. In the technology-rich classroom, a third source of information are the recordings of learning activities as they unfold in digital media, such as software applications (e.g. MS Excel), learning management systems (e.g. Moodle), and increasingly on "Cloud" tools and services (such as Google Docs). While these digitally enacted learning

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 556–559, 2013.

activities are easily recorded (e.g., as log files), they usually need further processing in order to be interpretable as information on learning and knowledge. So far, methods to do this automatically have been confined to so-called Intelligent Tutoring Systems and personalized learning systems [3]. This approach, however, needs a very detailed analysis of the knowledge/skill structures to be learned, and a very fine-grained learning application in order to trace the students' actions on a level that is appropriate for the diagnostic algorithms.

In the NEXT-TELL project, we want to contribute to classrooms where ICT is used to engage students in meaningful learning activities, and to provide teachers and students with nuanced information about learning when it is needed and in a format that is supportive of pedagogical decision making, thus optimizing the level of stimulation, challenge, and feedback density. The project philosophy is that teachers should not only be seen as the users of classroom technologies, and the recipients of information, but also as the innovators of technology-supported teaching and assessment practices, and as the creators of knowledge about students' learning. The main research questions are: (1) What (real time) data do teachers need for monitoring their students' learning activities and diagnosing their knowledge development? (2) How should these data be made accessible and represented (visualized)? (3) How can the data be gained from students (including methods for automated tracing)? (4) How can machine interpretation of students' learning activities and products be integrated with teachers' workflows?

2 An Open, Model-Based Approach to Teaching in 21st Century Classroom

Quality data on students' learning are not just "found" in the classroom or in log-files, but need to be explicitly and carefully produced. Teachers are trained to glean information from observing students' reactions and from inspecting their work, in addition to performing formal and informal assessments (exam, quizzes, etc.). For the ICT-rich classroom, many of these observational and assessment practices need to be modified, and the data that become additionally available need to be integrated into teachers' practices. To facilitate this, we build on a "glass-box" approach to technology-enhanced formative assessment; both the diagnostic data transformation procedures as well as the resulting learner model should in principle be open. And beyond that: not only should the assessment be 'open', it should also be the teacher who, in principle, can develop and modify technology enhanced classroom assessment methods. This takes also into account that diagnostic information on students' learning can come from many sources: we consider in particular teachers, students, (in the role of self and peer assessors), parents, and software applications. They all can produce diagnostically relevant information.

An important consideration in NEXT-TELL is that all assessment methods, independent of who employs them (teacher, student, parent, software) should adhere to certain quality criteria, in particular concerning their validity and reliability. For establishing validity, we build on the Evidence-centered Assessment Design methodology [4], and for establishing reliability of assessments in NEXT-TELL they are (largely automatically) recorded. In short, we require that teachers as well as computational assessment services describe how they, starting from observations on what learners do in the course of a learning activity (performance), and from the artifacts produced in the course of a learning activity, come to conclusions about learners' knowledge and skills.

NEXT-TELL equip teachers with an authoring tool for designing learning activity sequences and for relating these to expected knowledge changes (learning progressions) as well as assessment methods.

At the same time, students are supported in thinking about learning goals (and the relation between learning goals) as well as to identify a sequence of learning steps, including self- and peer-assessment steps.

In NEXT-TELL any assessment process is treated as an instantiation of an assessment model, and any learning activity sequence as an instance of a learning sequence model. Technically, we use a meta-modeling shell and the Open Models approach [5] for modeling formative assessment processes and learning activity sequences. The models are not only descriptions, but can also serve as the basis for rapid deployment in an ICT environment. Currently, we provide adaptors to Moodle, Mahara, and to Google Apps. We call the method and the toolkit ECAAD, for Evidence-centered Activity and Assessment Design.

Once an IT-based learning sequence is described in such a model (and as a side effect stored on a server), it becomes shareable (e.g., between teachers) and it constitutes a basis for technical deployment. Since the model describes the learning activities (e.g., those involved in peer writing) in some detail, the activities can be supported quickly in different software tools (e.g., Moodle, Google Apps).

And since the models specify the data that should be traced and/or requested from students during learning, it makes it easier to automatically gather and interpret log file information in terms of students' learning. Since such tracing data are sensitive due to trust and privacy issues, in NEXT-TELL all trace and log file data get stored in an e-Portfolio system (an extension of the open source tool Mahara), and are thus under student control. The student can control what to share in general, and what to share with the teacher for appraisal in particular.

Appraisals of students' work in NEXT-TELL are maintained in an Open Learner Model (OLM) [6], and made available to teachers and students in multiple modalities. The OLM represents students' proficiencies in terms of competences and standards. Entries in the OLM are not only "open for inspection", but also "open for negotiation" between stakeholders (teachers, students, possibly parents). A special tool for supporting evidentiary argumentation supports the negotiation part.

3 Conclusions

A study (financed by the Gates foundation) found that most school dropouts in the States today are due to boredom, not due to performing poorly [7]. In the future classroom, students will be less bored, but also less often over-challenged. Instead most students will work on their optimal level of cognitive density[1]; classrooms have become high-performance environments [2]. With the described methods and tools, the NEXT-TELL project aims to support teachers to diagnose what that optimal level is and monitor it for relevant sub-groups if not for each student individually. Teachers are dynamically adjusting their teaching based on real time data on their students' learning. For each student, they maintain an individual learning plan, agreed with the student and her parents. Students themselves are taking over significant parts of planning and monitoring their learning, guided by their teachers. The project is now finalizing the integration of the different tools developed with the assistance of teachers groups from Norway, Austria, UK and Germany. The last phase of the project will be dedicated to the implementation of several pilots with schools aiming at evaluating the actual effectiveness of our approach and tools.

Acknowledgements. This research has been co-funded by the European Commission within the FP7 ICT TEL project NEXT-TELL, www.next-tell.eu.

References

1. Hamilton, E., Jago, M.: Toward a theory of personalized learning communities. In: Jacobson, M.J., Reimann, P. (eds.) Designs for Learning Environments of the Future, pp. 263–282. Springer, New York (2010)
2. Crawford, V.M., Schlager, M.S., Penuel, W.R., Toyama, Y.: Supporting the art of teaching in a data-rich, high-performance learning environment. In: Mandinach, E.B., Honey, M. (eds.) Data-driven School Improvement, pp. 109–129. Teachers College Press, New York (2008)
3. Heller, J., Steiner, C., Hockemeyer, C., Albert, D.: Competence-based knowledge structures for personalised learning. International Journal on E-Learning 5(1), 75–88 (2006)
4. Bull, S., Kay, J.: Student models that invite the learner in: The SMILI() Open Learner Modelling Framework. International Journal of Artificial Intelligence in Education 17(2), 89–120 (2007)
5. Karagiannis, D., Grossmann, W., Hoefferer, P.: Open Model Initiative. A feasibility study. University of Vienna, Vienna (2008)
6. Brigdeland, J., Dilulio, J., Morison, K.: The silent epidemic: Perspectives of high school dropouts. Civic Enterprises, Washington, DC (2006)
7. Crawford, V.M., Schlager, M.S., Penuel, W.R., Toyama, Y.: Supporting the art of teaching in a data-rich, high-performance learning environment. In: Mandinach, E.B., Honey, M. (eds.) Data-driven School Improvement, pp. 109–129. Teachers College Press, New York (2008)

[1] Cognitive density describes "...the aggregate level of students' engagement with learning materials and thinking, their progress in learning, their communication and their use of time – that is, productive activity in the classroom at any given time" [7].

Signal Orchestration System for Face-to-Face Collaborative Learning Flows

Davinia Hernández-Leo, Raul Nieves, Juan Pablo Carrascal, and Josep Blat

ICT Department, Universitat Pompeu Fabra, Barcelona, Spain
{davinia.hernandez,raul.nieves,
juanp.carrascal,josep.blat}@upf.edu

Abstract. This demonstration paper is focused on the Signal Orchestration System (SOS), a system that helps teachers in the implementation of dynamic collaborative learning flows in physical spaces, such as the classroom. It is based on the use of color, haptic and sound signals to communicate changing orchestration indications of collaborative learning flows such group formation and distribution of activities, roles, collaboration areas, resources, etc. The system is composed of personal devices (worn by students), located devices (in specific physical areas), and a manager (installed in the teacher's computer to define and send the signals). Experimental results indicate that the SOS enables flexible orchestration of collaborative flows, facilitates orchestration awareness, decreases the time devoted to orchestration tasks and favors teachers' and students' attention to the learning task. The demonstration outline includes the on-the-fly configuration of the manager and the participation of attendees who, wearing the personal devices, are asked to react to the orchestration signals received in their personal devices, other attendees' devices and located devices in the room.

Keywords: CSCL, orchestration, physical spaces, wearable devices, roomware.

1 Background: Face-to-Face Collaborative Learning Flows

Collaborative Learning (CL) techniques rely on the role of social interaction as a mechanism potentially leading to effective learning outcomes. To increase the probability of achieving fruitful learning, CL flow patterns propose well-known dynamics to structure the group formation and the distribution of activities, roles and resources [1]. The rationale behind the flows proposed by the patterns focuses on facilitating situations of mutual explanation and regulation, socio-cognitive conflicts, positive interdependence, individual accountability, etc. For example, the Jigsaw pattern formulates a CL flow where a global problem is divided into sub-problems and distributed among team members. Then, the participants of the different teams (centered on the same sub-problem) form expert groups in order to exchange ideas about their sub-problem. Finally, the members of the original team meet again and each of them contributes with their sub-problem expertise to solve the global problem. Other CL flow patterns are the Pyramid or the Role-play [1].

However, the orchestration task of implementing these collaborative dynamics in face-to-face situations, i.e. in physical spaces such as the classroom, is often very

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 560–564, 2013.

demanding, especially when the flow dynamics are complex or the number of students is relatively high [2-3]. Teachers have to indicate group formation and role assignment for every activity, considering the use of multiple resources/tools (which may be allocated depending on the groups/roles), the distribution of spaces, and the supervision of the real-time learning progress. Both teachers and students need to devote part of their attention to orchestration aspects, which are also typically time-consuming and can generate a noise / mess effect leading to disorganization and distraction. The system described in this demonstration paper aims to tackle these problems. It is the Signal Orchestration System (SOS), a system based on the use of ambient displays to communicate orchestration signals flexibly configured in a manager, according to the teachers' needs and the characteristics of specific CL flows (*e.g.*, in a Jigsaw flow: color signals may indicate group formation and distribution of resources, sound or vibration signals change of activity, etc.).

The SOS is being defined following a design-based research methodology [4], in which iterative prototypes have been implemented and evaluated in real face-to-face settings with several CL flows [5]. While the first prototypes[1] explored how diverse designs for SOS personal wearable devices support better ergonomics and orchestration awareness [6], the iteration presented in this demonstration extends its features enabling the display of orchestration signals in specific locations of the space and the participation of students in shaping the orchestration.

2 Background: Signal Orchestration System (SOS)

Fig. 1 shows the current iteration of the SOS, which now includes located ambient displays and supports bidirectional communication. The prototype consists of three main components: the Personal Signaling Devices (PS-Device), the Located Signaling Devices (LS-Device) and the Signal Orchestration Manager (SO-Manager).

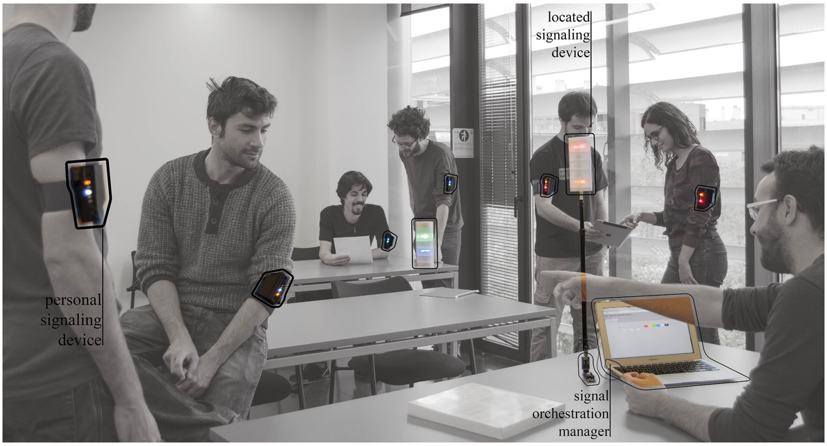


Fig. 1. Components of the current iteration of the Signal Orchestration System

[1] Video of prototype with the personal wearable devices supporting better ergonomics and orchestration awareness (vs. previous prototypes [5-6])
`http://www.youtube.com/watch?v=YtIYefbOs6A`

A PS-Device allows direct interaction on the participant's end. It is a small, physical device worn by every student (and other actors, if applies) participating in the collaborative activities. It includes a set of 5 LEDs with different colors, a vibrator (or buzzer) and two push-buttons. The LEDs and vibrator display orchestration signals to the students (by means of different combinations of colors, blinking colors, vibration), while the buttons allow them to send signals to the teacher (*e.g.*, indicating completion of activities, volunteering for roles…). Physically, the PS-device is a small black box that adheres to a band that wraps around the participant's arm, in an unobtrusive and comfortable way [6]. Inside, there is a microcontroller with a built-in transceiver (a Jeenode), which provides a two-way wireless communication with the OS-Manager, as well as a rechargeable battery pack [5]. An important feature of the PS-Devices is that the light signals are visible not only to the participant wearing the device, but also by the rest of the group, facilitating ambient orchestration awareness. This is one of the advantages of the SOS over a mobile phone based solution. Phones could display more detailed orchestration information, but in a less smooth, shared and persistent (also distraction-avoiding) manner. In any case, the SOS provides a minimalist ambient orchestration approach that can be used with other devices (phones, tablets, etc.) in a complementary way depending on the activities' requirements.

The LS-Devices were designed to be used as a "marking posts" for the physical space where the collaborative activity occurs. They help to identify and assign signals to specific places or shared physical resources (*e.g.*, collaboration areas, tabletops…), as opposed to the PS-Devices, which are meant for participants. Internally, the LS-Device uses the same hardware as the PS-Device, but its external appearance is different due to its specific goal. It has the shape of a vertically oriented cylinder that can be placed a table or another easy-to-spot surface, or also on top of a stand (with adjustable height), allowing the device to be positioned on the floor. The LEDs radiate to every side of the cylinder, dividing its surface in colored "bands", facilitating their visibility from every angle.

The OS-Manager is a software program running on a computer connected to another Jeenode [5], for wireless communication. It provides the teacher with a straightforward way to send and receive signals to and from the PS-Devices and LS-Devices. OS-Manager's functions include configuring custom signal assignation depending on the criteria required by the CL flow, and applying these assignations to students and locations. It allows editing assignation configuration on the fly, in order to flexibly address unexpected changes in the class, such as sudden absences or newcomers.

3 Results and Outcomes Achieved

Experimental results indicate that the SOS enables flexible agile orchestration of face-to-face collaborative learning flows, facilitates orchestration awareness, decreases the time devoted to orchestration tasks and favors teachers' and students' attention to the learning task [5]. Three different designs for the PS-Devices have been implemented and tested in real settings: a necklace device was more visible, but its size and weight made it more uncomfortable; a fabric belt was lighter, thinner and aesthetically nicer, but it was less visible; a compact arm bracelet (Fig. 1) as an intermediate approach whose position in the arm facilitates the visibility of the signals even when the participants are sitting down at their desks [6]. When compared with a control group (results reported in a paper currently under review), students using the PS-devices showed statistically significant improvements in terms of reduction of time devoted to orchestration tasks, perception of group awareness, and performance.

New experiments focused on the effects of the LS-devices are currently being designed. Moreover, the system is being integrated with virtual learning environments to enable a seamless transition between activities in digital and physical spaces in the same CL flow (*e.g.*, in a Jigsaw flow: expert group formation in a Moodle environment according to the sub-problem distribution realized in a previous activity in the physical classroom). All in all, towards the ultimate aim of scaling up and sustaining in practice the implementation of orchestrated collaborative learning flows.

4 Demonstration Outline

The planned demonstration consists in a collaborative flow involving the participation of the conference attendees. The demonstration will be focused on showing the added value – in terms of features - of the SOS: (1) remote configuration of ambient displays (wearable devices, devices located in zones or furniture) with orchestration indications; (2) pre-planning of the collaborative learning flow to be carried out face-to-face in the physical space and fast deployment; (3) real-time modifications on the fly, without interrupting the flow; (4) digitalized information of the orchestrated flow enables integration with other systems, such as virtual learning environments, mobile phone applications, etc. The demonstration will also point out experimental results indicating that the SOS minimizes classroom "chaos" and the time devoted to orchestration while favoring orchestration awareness, enjoyment and flexibility.

Acknowledgments. This work has been partially funded by the Spanish Ministry of Economy and Competitiveness, EEE project TIN2011-28308-C03-03.

References

1. Hernández-Leo, D., Asensio, J.I., Dimitriadis, Y., Villasclaras, E.D.: Pattern languages for generating CSCL scripts: from a conceptual model to the design of a real situation. In: Goodyear, P., Retalis, S. (eds.) E-learning, Design Patterns and Languages, pp. 49–64. Sense Publishers, Rotterdam (2010)
2. Mäkitalo-Siegl, K., Zottmann, J., Kaplan, F., Fischer, F.: Classroom of the Future: Orchestrating Collaborative Spaces. Sense Publishers, Rotterdam (2010)
3. Dillenbourg, P., Jermann, P.: Technology for classroom orchestration. In: Khine, I.M., Saleh, I.M. (eds.) The New Science of Learning, pp. 525–552. Springer, New York (2010)

4. Amiel, T., Reeves, T.C.: Design-based research and educational technology: rethinking technology and the research agenda. Educational Technology & Society 11(4), 29–40 (2008)
5. Hernández-Leo, D., Nieves, R., Arroyo, E., Rosales, A., Melero, J., Blat, J.: SOS: Orchestrating Collaborative Activities Across Digital and Physical Spaces Using Wearable Signaling Devices. Journal of Universal Computer Science 18(15), 2165–2186 (2012)
6. Hernández-Leo, D., Balestrini, M., Nieves, R., Blat, J.: Exploiting awareness to facilitate the orchestration of collaborative activities in physical spaces. In: 2nd Workshop on Awareness and Reflection in Technology-Enhanced Learning, EC-TEL2012. CEUR WS, vol. 931, pp. 55–56 (2012)

GLUEPS-AR: A System for the Orchestration of Learning Situations across Spaces Using Augmented Reality

Juan A. Muñoz-Cristóbal, Luis P. Prieto, Juan I. Asensio-Pérez,
Iván M. Jorrín-Abellán, Alejandra Martínez-Monés, and Yannis Dimitriadis

GSIC-EMIC Group, Universidad de Valladolid, Spain
{juanmunoz,lprisan}@gsic.uva.es,
{juaase,yannis}@tel.uva.es, ivanjo@pdg.uva.es,
amartine@infor.uva.es

Abstract. Ubiquitous and mobile learning scenarios define activities happening within and beyond the walls of a classroom. However, the orchestration of authentic learning situations involving both physical and virtual spaces is still a significant challenge for teachers. Several proposals recently reported in the literature try to reduce teachers' orchestration burden by means of authoring tools, and usage of Augmented Reality technologies for connecting physical and virtual spaces. However, these proposals are restricted to specific technologies, pedagogies, or to a very limited range of activities. We present GLUEPS-AR, a system to aid teachers in the orchestration of across-spaces learning situations. With GLUEPS-AR, learning designs defined with multiple authoring tools can be deployed and managed at run-time throughout ubiquitous learning environments composed by different VLEs, Web 2.0 tools and AR applications. Thus, GLUEPS-AR allows multiple design and enactment technologies, and a wide range of learning activities, not restricted to a single pedagogy.

Keywords: Augmented reality, across spaces, ubiquitous learning, VLE, orchestration.

1 Background (Pedagogy)

Different pedagogical approaches explore learning beyond the traditional classroom, such as ubiquitous, situated or place-based learning. A fundamental issue in these ubiquitous contexts is to maintain a continuous learning experience across the different spaces where learning occurs [1]. This is especially challenging in collaborative scenarios, which pose still more technology and pedagogical difficulties. Augmented reality (AR) is a technology that may be used to connect physical and virtual spaces, as AR is based on the superposition of virtual objects in physical environments [2]. AR has also pedagogical affordances, such as the capability of enhancing spatial skills using 3D virtual objects, the potential positive impact on the engagement of students, and the scaffolding in learning (e.g., guiding students with AR).

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 565–568, 2013.

2 Background (Technology)

Existing research work provides a variety of solutions for the deployment of across-spaces learning situations using AR to connect physical and virtual spaces. Interestingly, these initiatives typically use learning design processes to facilitate teachers' deployments, in the form of authoring tools [3,4]. Learning design [5] is a field of research in technology enhanced learning, based in the explicit formulation by practitioners of their pedagogical ideas, which are often defined in a computational form.

However, the existing across-spaces prototypes have limitations that may complicate their adoption by teachers: a) these prototypes work with a limited range of authoring and enactment tools (usually, only one of each type); b) some of these prototypes are also oriented to specific pedagogies (e.g., game based learning); and c) they are usually restricted to a particular kind of learning activities (e.g., activities consisting on a sequence of pre-specified AR artifacts). However, teachers need more flexible approaches, able to accommodate to the technologies and pedagogies that they already use in their daily practice.

3 Results and Outcomes Achieved

We propose GLUEPS-AR [6] to aid teachers in the orchestration of across-spaces learning situations. GLUEPS-AR is a system for the deployment of learning designs (which may be defined using multiple authoring tools) into multiple ubiquitous learning environments. Such ubiquitous environments are composed by Virtual Learning Environments (VLE), AR-enabled physical spaces, and Web 2.0 tools. GLUEPS-AR also allows the run-time management of such learning situations.

We have defined a technological architecture that extends a predecessor system centered on distributed Web spaces called GLUE!-PS[1]. The architecture is based on a number of adapters that enable the integration of authoring tools (e.g., WebCollage[2] or Pedagogical Pattern Collector[3]), VLEs (e.g., Moodle[4]), AR applications (e.g., QR code readers or Junaio[5]) and Web 2.0 tools (using multi-tool managers, such as GLUE![6] or IMS-LTI[7] to integrate multiple external tools in VLEs).

By means of these adapters, Web 2.0 tools and other virtual artifacts (e.g., 3D models) may be accessed from both VLEs and physical spaces, in a global across-spaces learning situation. Virtual learning artifacts may be positioned in physical spaces through AR with different positioning types (e.g., QR codes, visual markers or

---

[1] `https://www.gsic.uva.es/glueps/`. Last access, April 2013.
[2] `http://pandora.tel.uva.es/wic2/`. Last access, April 2013.
[3] `http://thor.dcs.bbk.ac.uk/projects/LDSE/Dejan/ODC/ODC.html`. Last access, April 2013.
[4] `http://moodle.org`. Last access, April 2013.
[5] `http://www.junaio.com`. Last access, April 2013
[6] `https://www.gsic.uva.es/glue/`. Last access, April 2013.
[7] `http://www.imsglobal.org/toolsinteroperability2.cfm`. Last access, April 2013.

geo-position). As GLUEPS-AR is not restricted to a reduced number of design or enactment technologies, it may support multiple pedagogies and activity types.

A prototype has been developed to illustrate the architecture and test it. Fig. 1 shows GLUEPS-AR's architecture and user interface. The interface allows to modify, complete and configure the initial design (e.g., including concrete participants/groups, adding tool configurations, etc.). The Figure also displays the different authoring tools and learning environments (VLEs and AR applications) supported by the prototype. We have accomplished an initial evaluation, co-designing a complex across-spaces situation with a faculty member of the College of Education at the Universidad de Valladolid (Spain). The teacher co-deployed the learning situation using GLUEPS-AR and simulated part of the enactment, giving positive feedback about the deployment process and the possibilities of the approach. This learning situation, as well as others currently being designed, are planned to be realized with actual students in the upcoming academic year.

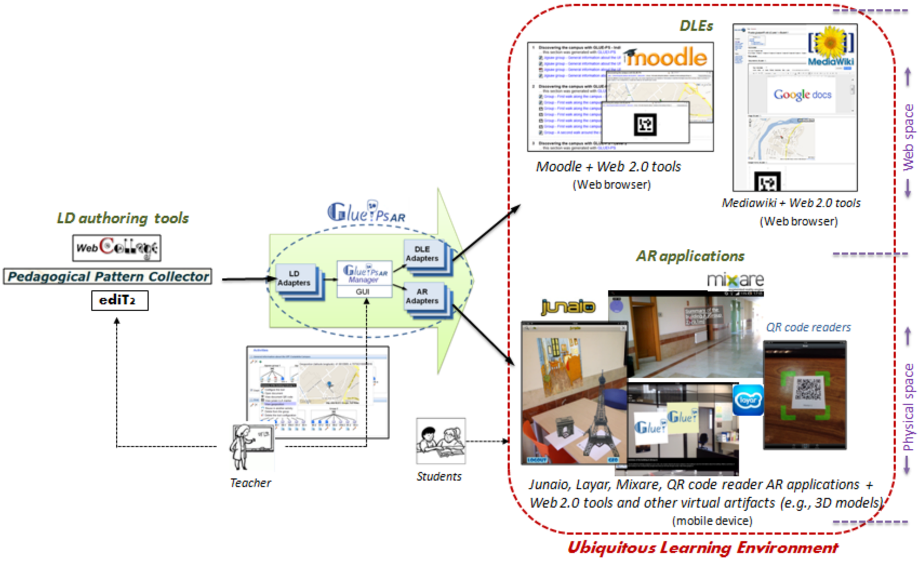


Fig. 1. GLUEPS-AR architecture (center), and design and enactment tools currently supported by the prototype (left and right, respectively)

4 Demonstration Outline

The demonstration will show how an initial across-spaces learning design (that has been previously defined in an authoring tool, such as WebCollage) can be deployed in minutes into a ubiquitous learning environment composed by a VLE, Web 2.0 tools (Web space), and an AR application (physical space). The setting up of the learning design ready for enactment in the ubiquitous learning environment can be verified with computers and mobile devices (tablet and smartphones) simulating how students would access the learning artifacts available in different spaces during enactment.

5 Specific Technology and Environment Needed at Conference

Demonstration will be performed by 1-2 demonstrators. A good wireless (Wifi) Internet connection is required, so that mobile devices can download AR objects to be rendered. We will use laptops, tablets and smartphones (provided by us). We would need a projector and VGA cable (to connect projector and laptop), and a surface for projection (an alternative to the projector could be a PC monitor with at least 21'). If there is a demonstration session where people cannot be close to our laptop and mobile devices, (e.g., a plenary session with a sequence of demonstrations towards a sat audience), we would need two projectors (in order to project simultaneously laptop and tablet screens), and two VGA cables.

Acknowledgements. This research has been partially funded by the Spanish Projects TIN2011-28308-C03-02 and VA293A11-2, and the European Project 531262-LLP-2012-ES-KA3-KA3MP. The authors thank as well the rest of the GSIC/EMIC research team.

References

1. Sharples, M., Sanchez, I.A., Milrad, M., Vavoula, G.: Mobile learning: Small devices, big issues. In: Balacheff, N., Ludvigsen, S., de Jong, T., Barnes, S. (eds.) Technology Enhanced Learning: Principles and Products, pp. 233–249. Springer, Heidelberg (2009)
2. Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent advances in augmented reality. IEEE Computer Graphics and Applications 21(6), 34–47 (2001)
3. Klopfer, E., Sheldon, J., Perry, J., Rosenheck, L., Squire, K., Mathews, J., Shapiro, R.B., Coulter, B., Dunleavy, M.: Augmented reality games: Place-based digital learning. In: Proceedings of the 9th International CSCL Conference, Hong Kong, China (July 2011)
4. Santos, P., Pérez-Sanagustín, M., Hernández-Leo, D., Blat, J.: QuesTInSitu: From tests to routes for assessment in situ activities. Computers & Education 57(4), 2517–2534 (2011)
5. Koper, R.: An Introduction to Learning Design. In: Koper, R., Tattersall, C. (eds.) Learning Design Modelling and Implementing Networkbased Education Training, vol. 2003, pp. 3–20. Springer (2005)
6. Muñoz-Cristóbal, J.A., Prieto, L.P., Asensio-Pérez, J.I., Jorrín-Abellán, I.M., Dimitriadis, Y.: Orchestrating TEL situations across spaces using Augmented Reality through GLUE!-PS AR. Bulletin of the Technical Committee on Learning Technology 14(4), 14–16 (2012)

Virtual Afghan Village as a Low-Cost Environment for Training Cultural Awareness in a Military Context

Ekaterina Prasolova-Førland[1], Mikhail Fominykh[1], and Ramin Darisiro[2]

[1] Program for learning with ICT, Norwegian University of Science and Technology
[2] Norwegian Armed Forces
{mikhail.fominykh,ekaterip}@ntnu.no, rdarisiro@fhs.mil.no

Abstract. While 3D virtual worlds, simulations, and serious games have been used for military training for a long time, there is a lack of research-based methods for using and creating such systems. In addition, the development process has traditionally been very resource consuming. The project Cultural Awareness in Military Operations seeks to address these challenges. The goal of the project has been to create and evaluate a low-cost, off the shelf, game-based simulation for training cultural awareness among military personnel in the Norwegian Armed Forces preparing for international operations. The project has also aimed at creating methodological guidelines and tools for developing 3D educational simulations for both military and civilian use.

Keywords: Military training, 3D virtual worlds, simulations, cultural awareness.

1 Background (Pedagogy)

In this paper, we present the results of the project Cultural Awareness in Military Operations (CAMO). Its goal has been to create an inexpensive and flexible game-based simulation for training cultural awareness among military personnel in the Norwegian Armed Forces preparing for international operations (Afghanistan). The project is a joint effort of the Norwegian Armed Forces, the Norwegian University of Science and Technology, and the University of Oslo.

Cultural awareness in this context or *operational culture* is defined as "those aspects of culture that influence the outcome of a military operation; conversely, the military actions that influence the culture of an area of operations" [1]. Studying culture allows (theoretically) military units and partners to understand specific actions of locals, exert influence on the population, and improve interaction with other players in the area of operation. Practically applicable knowledge was conveyed through practical learning objectives that had been embedded into the training including a virtual Afghan village and local 'Afghans' in a virtual world (VW) built in Second Life (SL). The project also aimed at creating research-based methodology, guidelines, and tools for developing 3D educational simulations for the Norwegian Armed Forces [2]. This was motivated by an identified lack of research-based methods for using game-based simulations in military training, both within the Nordic armed forces and among other allies (NATO) [3].

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 569–572, 2013.

As a starting point for developing such a methodology, we studied publicly available literature and guidelines for operational culture training [1] and scenario methodologies, see e.g., [4]. These methodologies focus primarily on tactical tasks and not on operational culture, they have been developed further to be adjusted to the goals of the CAMO project [2]. Since the existing literature on Afghani operational culture is rather fragmented and/or classified, we used subject matter experts within the Norwegian Armed Forces and Norwegian academia as the major source of information.

The methodology for scenario development used for the CAMO project is based on a systematized set of learning goals and associated 'mini-scenarios', to obtain the maximum reusability of the content. The following major categories of learning goals were identified: *Tactics*, *Gender*, *Religion*, *Socializing* and *Language*. Each of the learning goals categories are split into sub-categories, providing a basis for the corresponding 'mini-scenarios'. Each of the learning sub-goals is further detailed with corresponding cues, appropriate reactions, typical mistakes, and typical responses in case of mistake. Though initially developed for the project setting (focusing on international operations in Afghanistan), this methodology is generally applicable for operational culture/cultural awareness training of both military and civilian personnel.

2 Background (Technology)

3D VWs and game-based simulations have been used for military training for a long time, demonstrating concepts and situations that are difficult, expensive or unsafe to represent efficiently in a classroom, including culture [5]. Examples of such systems for training operational culture include Tactical Iraqi and First Person Cultural Trainer [6,7]. They are typically very expensive to develop and primarily single-player, providing no or very limited support for collaborative learning and team training, with very few possibilities to modify and generate new scenarios.

Scenarios for the CAMO project have been developed according to the methodology mentioned in the previous section, providing requirements for the design and implementation of the virtual environment. Later, due to various constraints, trade-offs and re-focusing of the game's learning objectives, both the scenarios and environment design needed to undergo certain adjustments and simplifications.

Our focus during implementation was on low cost, short development time, and reusability. The overall project scenario describes the environment for the educational simulation – the Afghan village, while each mini-scenario requires additional specific content. The environment consists of the general content for creating the context (e.g., landscape elements, vehicles, furniture, authentic clothing) and specific content for mini-scenarios (a mosque, a school, a medical kit, a camera, a specific gun, and tableware). The implementation of the environment went through several stages. First, the required objects had to be created or collected. Practically, some of the objects and avatar clothing have been acquired from the SL marketplace as well as searching free objects everywhere in SL. However, most of the specific content artifacts were designed from scratch to ensure authenticity. Second, when the basic objects (or elements) were collected and platform was set up, building and co-locating the typical elements started, reusing them in multiple places of the environment or joining in different combinations. After receiving feedback from the subject matter experts, some of them were modified or replaced.

3 Results and Outcomes Achieved

The simulation was evaluated in an experiment conducted at the Norwegian Military Academy on November 25, 2011. It was preceded by a 'rehearsal' session the week before and two-hour introduction training in SL technology. The participants included 14 cadets from the Norwegian Military Academy, playing roles of the Norwegian soldiers in the simulation (Figs. 1, 2) as well as six students and two teachers from the Norwegian Defense Language and Intelligence School, playing the roles of the Afghan civilians and interpreters and providing input to the scenario development.

Fig. 1. Conversation with the village chieftain

Fig. 2. Contact with local women

The role-play was organized in two rounds, including three debrief sessions conducted by an expert in Afghan culture. In each of the rounds, a group of the cadets/ 'squad' executed the mission, proceeding through the different scenes of the scenario. In the first round, the squad leader had a previous field experience from Afghanistan.

The data have been collected through observation of the role-play, screen-capture recording, pre- and post-questionnaires, and in-depth interviews of selected participants. The interview data provided further indications that the soldier' understanding of cross-cultural issues has improved over the course of the experiment. For example, one of the female participants noted: "I got very much out of it during a very short time", "plenty of aha-experiences". She also reported a high level of immersion in her role. Another, male participant believed that: "This (system) can provide several possibilities in a deployment environment to increase understanding among troops preparing for international operations".

It was not practically possible to access the participants' performance during potential deployment, but the questionnaire data give tentative evidence of an increased understanding of the cultural and religious aspects. At the same time, some of the soldiers disagreed that the experience in SL was suitable for correct evaluation of the threat situation due to the lack of 'crowd', which could possibly indicate an ambush. The cadets have been generally positive to the use of 3D VWs for training cultural awareness. On the overall, they reported that the simulation in SL has been a user-friendly, motivating, and fun experience. At the same time, the participants identified a number of limitations, especially, a limited selection of avatar gestures and body language that complicated expression and perception of certain cultural and social aspects. In addition, not all the players have been equally engaged in the game, something that affected their learning outcomes.

Following the feedbacks, the virtual Afghan village will be improved. It is planned to develop agents for crowd simulation of a 'regular' village life. Additional scenarios will be developed to account for a greater variety of learning goals and including both military and civilian personnel. At the same time, it is not possible to improve certain aspects such as limited body language and mimics range (inherent to SL) without significant increase of the cost. It is therefore necessary to consider a trade-off between the cost and quality in relation to different learning goals.

Acknowledgements. The authors wish to thank Øystein Ramseng, David Hansen, Ingvill Thomassen, and Anders Mørch for contributions, as well as study participants.

References

1. Salmoni, B.A., Holmes-Eber, P.: Operational Culture for the Warfighter: Principles and Applications. Marine Corps University Press, Quantico (2008)
2. Prasolova-Førland, E., Fominykh, M., Darisiro, R., Mørch, A.I.: Training Cultural Awareness in Military Operations in a Virtual Afghan Village: A Methodology for Scenario Development. In: 46th Hawaii International Conference on System Sciences (HICSS), Wailea, HI, USA, January 7–10, pp. 903–912. IEEE (2013)
3. Singer, M.J., Knerr, B.W.: Evaluation of a Game-Based Simulation During Distributed Exercises. U.S. Army Research Institute for the Behavioral and Social Sciences, Research Report 1931 (2010)
4. Hartog, C.: Scenario design for serious gaming: guiding principles for the design of scenarios and curricula in military Job Oriented Training. TNO Defense, Security & Safety (2009)
5. Ogan, A., Lane, H.C.: Virtual Learning Environments for Culture and Intercultural Competence. In: Blanchard, E.G., Allard, D. (eds.) Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models, pp. 501–519. IGI Global, Hershey (2010)
6. Zielke, M.A.: The First Person Cultural Trainer Whitepaper (2011), http://www.utdallas.edu/~maz031000/res/FPCT_White_Paper.pdf
7. Surface, E.A., Dierdorff, E.C., Watson, A.M.: Special Operations Language Training Software Measurement of Effectiveness Study: Tactical Iraqi Study Final Report. Special Operations Forces Language Office, Tampa, FL, USA (2007)

etiquetAR: Tagging Learning Experiences

Mar Pérez-Sanagustín, Alejandro Martínez, and Carlos Delgado Kloos

Departamento de Ingeniería Telemática,
Universidad Carlos III de Madrid, Av. Universidad, 30.
Edif. Torres Quevedo, 28911 Leganés, Madrid
{mmpsanag,alemarti,ckd}@it.uc3m.es

Abstract. etiquetAR is an authoring tool for supporting the design and enactment of mobile context-based learning experiences based on QR tags. etiquetAR enables creating, managing, and sharing personalized QR tags attachable to any object or physical geographical location. Tags are digital layers of contextualized information that transforms any physical space into a digitally augmented learning environment. This demonstration paper presents etiquetAR first working prototype of this application. In particular, the paper details: (1) how etiquetAR web-based application can be used to edit a tag, associate different resources to it, and relate resources information to a particular profile for adaptive learning experiences; and (2) how users can access and contribute to the information hidden in the tags using the mobile-based application. This demonstration will show the audience how etiquetAR is a simple tool designed to encourage practitioners to create their own tag-based learning experiences.

Keywords: mobile learning, QR tags, augmented reality, mobile-web-based application, demonstration.

1 Pedagogical Background

etiquetAR[1] [2] is an application designed towards educational purposes based on mobile learning pedagogical theories. Mobile learning is a recent learning theory that "recognizes the essential role of mobility and communication in the process of learning and the importance of the context in establishing meaning construction" [2].

Sharples et al. (2010) [2] define Mobile learning as "the processes of gaining knowledge through conversations across multiple context amongst people and personal interactive technologies" and propose a framework to structure it. This framework expands the Engestrom's activity model [3] to tackle the interdependencies between learning and technology. The framework comprises five factors (subject, object, context, tools and communication) analysed under two perspectives or layers: the technological and the semiotic. The technological layer depicts learning as an engagement with technology, in which tools such as computers and mobile phones function as interactive agents in the process of gaining knowledge, creating a human-technology system with which to communicate, mediate agreements between learners

[1] http://etiquetar.com.es

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 573–576, 2013.

(as with spreadsheets, tables, and concept maps), aid recall, and reflection (as with weblogs and online discussion lists). The semiotic layer describes learning as a cognitive system in which learners' objective-oriented actions (i.e. actions that promote an objective) are mediated by cultural tools and signs. In this framework, technology acts as the facilitator.

etiquetAR was designed according to this framework for supporting knowledge construction in context. Specifically, etiquetAR is based on the idea that digital tags, when attached to a particular object or location, add a layer of digital information that extends its properties. Then objects and locations become interactive means through which learners, using their mobile personal devices, gain knowledge about their environment. In this context, tags act as facilitators. Through interaction with tags created with etiquetAR learners can reflect about a particular object or location, contribute with messages to these tags, and share these messages with others to construct and transform their learning contexts on the move. Therefore, etiquetAR is a tool that facilitates the design of activities that enhance situated, experiential and collaborative learning at the same time that promotes active learning and reflection in context.

2 Technological Background

The first prototype of etiquetAR [1] is based on tags that follow the QR standard [4]. QR standard is used to codify any information in a matrix that can be accessed using a reader following that standard. QR tags are typically used to codify an URL linking digital information in the cloud. In etiquetAR, QR tags are used as mechanisms to contextualize this digital information in a particular physical object or location, linking both digital and real worlds

etiquetAR is composed of two parts: (1) a web application that enable users to create, personalize, and manage QR interactive tags, and (2) a mobile application that benefits from any third-party QR reader conform with the standard for redirecting users to the information hidden in the tags organized into the different available profiles. The whole application has been developed using the Ruby on Rails framework, with a PostgreSQL database on the backend and designed to be run on the Heroku platform.

Currently there are many tools for creating QR tags from a plain text, an URL, or a contact card. However, none of these tools is educationally oriented. Moreover, existing tools do not take advantage of all the potentiality that QR tags offer since they do not cover the complete life-cycle of the tag: creating the tag, editing the content and updating that context. etiquetAR supports the whole life-cycle of QR tags. Moreover, etiquetAR has been designed towards educational purposes and has the following added values, thus differentiating itself from other QR codes generators:

- **Support for Creating Dynamic QR Tags:** etiquetAR enables creating QR tags with a fixed image and vary the content whenever the user wants.
- **Support for Creating Tags Linking to Multiple Digital Resources:** etiquetAR enables to create more than one resource for the same tag and to manage its access through profiles in order to adapt the learning process to any type of learner.

- **Support for Creating Adaptive Tags:** etiquetAR enables assigning a profile to access the codified information. Each resource in the tag can be assigned to a different profile. This profile determines which resource is shown to the user.
- **Support for Organizing Tags into Collections:** Tags can be organized into folders, making easier for users to access their tags (i. e. Collection "Moma Museum", "Botanic garden", etc.). One tag can be associated to one or more folders.
- **The Tags Can Be Accessed with Any QR Reader:** Since the tags follow the QR standard, any readers conformed to this standard can read the tags created with etiquetAR.
- **Mobile-Based Support for Contributing to Existing Tags:** Any user can add comments to the resources codified in a tag generated with etiquetAR. Also, any user can see the comments provided by others to a particular resource. The owner of the QR codes can delete and hide any comment.

3 Results and Outcomes Achieved

etiquetAR has been used in a tag-based activity in a formal learning University context: "Discovering the Campus 2012". In this activity the tags generated with etiquetAR were used to augment the University Campus and support the students from the first course to learn about the Campus' services and areas. The students used their mobile devices to access to the information hidden in the different tags distributed around the Campus. Each tag was enriched with three resources associated to one of the courses participating in the activity through profiles: informatics, telematics and audio-visual systems undergraduate students[2].

4 Demonstration

The aim of this demonstration paper is to describe how etiquetAR web and mobile-based applications work. An illustrative scenario of a science activity in a museum will motivate the demonstration.

Two practitioners want to design a visit to a science museum. The objective of this activity is to make students reflect about the exhibits during their visit and to relate them to the content worked in class. The teachers use etiquetAR to prepare the content for the tags attached to the artwork. First, they login to the system and select the science museum collection. Then, they edit the existing tags to add resources (URLs, images, videos or text) and questions related to the content worked in class. They decide that each tag will contain resources related with three subjects: math, history, and science. For this purpose, they create three profiles, one for each subject. Then, they relate the different resources added previously to the tags to one of these profiles.

[2] The results of this activity are reported in a poster of the EC-TEL 2013 conference [5] as a complement of this demonstration paper.

Fig. 1. Web and mobile applications of etiquetAR

Once at the museum, students use their smartphones for reading the tags in the exhibits. They can use any QR reader installed in their devices. When accessing a tag, students are redirected to the profile selector for choosing the resource that they prefer to see. If the resource is a question, they can add their answer using the comment functionality and read the answers by their colleagues to complement them. During the visit, the teachers can use their smartphones to see the comments added to the different tags and even hide an eventual inappropriate comment.

Acknowledgements. This work has been partially funded by the Spanish Ministry of Economy and Competitiveness with the EEE project (TIN2011-28308-C03-01 and TIN2011-28308-C03-03) and by the eMadrid project (S2009/TIC-1650) funded by the Regional Government of Madrid. The authors would also like to especially thank the members of the research groups GAST (Universidad Carlos III de Madrid) for their contributions and ideas.

References

1. Pérez-Sanagustín, M., Martínez, A., Delgado Kloos, C.: EtiquetAR: A Tool for Designing Tag-based Mobile Augmented Reality Experiences. IEEE Technical Committee on Learning Technology 14(4), 27–30 (2012)
2. Sharples, M., Taylor, J., Vavoula, G.: Towards a theory of mobile learning. In: Proceedings of MLearn 2005, pp. 1–9 (2005)
3. Engestrom, Y.: Learning by expanding. An activity-theoretical approach to developmental research. Konsultit Oy, Orienta (1978)
4. QR Standard ISO, `http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43655` (retrieved March 29)
5. Pérez-Sanagustín, M., Melero, J., Hernández-Leo, D., Delgado Kloos, C.: Transforming the campus in a digitally augmented space using etiquetAR. In: Hernández-Leo, D., Ley, T., Klamma, R., Papadopoulos, G. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 625–626 Springer, Heidelberg (2013)

Tracking a Collaborative Web2.0 E-Learning Environment

Bénédicte Talon[1], Insaf Kerkeni[1], Sabra Tliche[2], and Henda Belaid-Ajroud[2]

[1] Univ Lille Nord de France, F-59000 Lille, LISIC, ULCO, F-62228 Calais, France
{Benedicte.Talon,Insaf.kerkeni}@univ-littoral.fr
[2] LIPAH, Faculté des Sciences de Tunis, Le Belvédère, Tunisie
tliche.sabra@gmail.com,
henda.ajroud@fst.rnu.tn

Abstract. In this paper, we present a multi-agent system that has been developed to trace activities in a collaborative e-learning environment. These traces are necessary to help the teacher to monitor the learners' activities and to allow the learners to identify their participation in the collective work. We also need traces to assist in the evaluation of the device. Indeed, knowing how the system has been really used facilitates its revision. The tracking system is implemented as a multi-agents system.

Keywords: E-Learn2.0 environment, trace, indicator, multi-agents system.

1 Introduction

In this paper, we present a system which tracks activities in collaborative e-learning environments. Among computer-supported learning environments exist E-Learn 2.0 environments. These environments combine Web 2.0 tools and make use of the different functionalities offered by these tools to produce an original education system. The advantages of the E-Learn 2.0 environments are presented in [1]. Here, we are interested in collaborative environment because it offers opportunities to develop knowledge and competences [2]. However, in an education system, the teachers should assess the work and/or consequently knowledge and competences acquisition. Moreover, teachers must insure that learners are active and are not out of play. In a face-to-face environment, the teacher has the ability to observe the learners' activity but in an on-line environment, the observation is more difficult. It is consequently important to propose a system for the activities observation or tracking. This system is in charge to collect the traces of the learners' activities and to produce information synthesizing their activities and enabling to identify the participation of each learner in a collaborative work.

In this paper, we first review the terminology related to traceability. We then present our specific context. Then we present a tracking tool developed for a collaborative web 2.0 learning platform. We conclude on the opened perspectives.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 577–580, 2013.

2 Trace, Indicator and Existing Systems – State of the Art

Several definition of a trace has been defined in the literature [3, 4]. These definitions depend on the role and the use of the trace and have small differences. In our TEL (Technology Enhanced Learning) context, we can retain that a trace may be a representation of activity, of use or of interaction within the learning system to which we add the temporal dimension. This trace must go through a set of transformations to lead to indicators which provide the user with a useful interpretation of the learning situation.

An indicator is a digital, alphanumerical or qualitative variable [4]. In the context of collaborative TEL, the indicators should compile traces to evaluate the learner individually and with his/her group, to evaluate acquired knowledge during the learning activity, to interpret the interaction between learners, to regulate the pedagogical device proposed by the teacher during a learning activity and to bring to the light some shortcomings of it and to customize the learning activity [5].

The tracking systems are concerned with the interpretation and the visualization of traces and indicators. A lot of tracking systems exist such as SPLACH [6], gStudy [7], Tatiana [8], TrAVis [9], IDLS [10] and SBT-IM [11]. However, each of these tracking systems has a different target, is specific to a particular platform and can't be used with a Web 2.0 learning platform.

3 Our Specific Context

We are interested in Web 2.0 collaborative learning environments. Learners are divided into groups and each group must do the work required by the teacher. The Web 2.0 learning platform provides various tools: blog, e-mail, forum and chat. The teacher posts the required work and the pedagogical devices on his/her blog. Each group of learners has its own blog on which deliverables are posted.

To enable the teacher assess the knowledge acquired by the learners, evaluate the state of the collaboration between the learners of the same group and know the usefulness of the pedagogical devices proposed to his/her learners, it is essential to observe during the learning activity all the actions done by the learners via the collaborative platform. Tracking learners and groups generates traces used to build indicators. These indicators allow the teacher to build the general assessment of the learning activity and allow the learner to do a self assessment. We define needed indicators according to the CAS taxonomy [12]: cognitive, activity and social indicators.

4 The Tracking System

We present succinctly the proposed tracking system. Figure 1 illustrates the functioning of the proposed system. The first step is the tracing step. Users' actions are saved in log files when they use the e-learning platform. These traces are modeled (second step) and directly stored (third step) in the traces database and then, the log files already treated are deleted. At this level, the teacher can request the database to view modeled traces in tabular form. The modeled traces are also used to compute indicators which will be available to be viewed by the appropriate actor (teacher or learner).

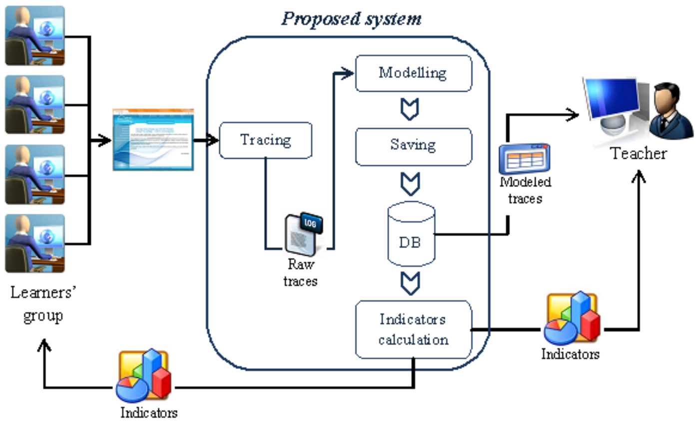


Fig. 1. Functioning of the proposed tracking system

The tracking system is a multi-agent system. This is motivated by several reasons as presented in [13]. To model the multi-agent system, we used MaSE [14] methodology. The tracking system includes the following agents:

- Tracking agents that supervise activities with e-learning collaborative tools. An agent is defined for each used web 2.0 tool. After modeling traces, these agents calculate indicators.
- Assistance agents that allow the presentation of information traced and analyzed by tracking agents to enable teachers to carry out the assessment of the activity and to provide assistance to learners.

The tracking system has been implemented using Madkit [15] platform and the Java programming language. The modeled traces obtained are stored in a MySQL database.

5 Conclusion and Perspectives

In this work, we are interested in the extraction of traces from a collaborative e-learning platform using Web 2.0 tools. These traces are modeled, stored and then processed to generate indicators. These indicators concern learner or group activities. They are used to support the evaluation of learners, groups, learning situations and pedagogical devices. On the other hand, the learner can do self-assessments.

The developed tracking system is a multi-agent system. It was tested within ILIAS platform [16] and it is planned to test it in real e-learning situations at "Institut Universitaire Technologique" (or IUT) of Calais. We also intend to upgrade the system by adding an intelligence level to the agents so that they will be able to make some decisions or to propose some decisions to the teacher.

References

1. Grosseck, G.: To use or not to use web 2.0 in higher education? Procedia Social and Behavioral Sciences 1, 478–482 (2009)
2. Honebein, P.C.: Seven goals for the design of constructivist learning environments. In: Constructivist Learning Environments: Case Studies in Instructional Design, pp. 11–24 (1996)
3. Georgeon, O.L., Mille, A., Bellet, T., Mathern, B., Ritter, F.E.: Supporting activity modelling from activity traces. Expert Systems 29(3), 261–275 (2012)
4. Djouad, T., Benmohammed, M., Mille, A.: Indicators computation from modeled traces in the context of computer Human Learning environment. In: International Symposium on Modelling and Implementation of Complex Systems, Constantine, Algérie, pp. 1–7 (2010)
5. Lund, K., Mille, A.: Traces, traces d'interactions, traces d'apprentissages: définitions, modèles informatiques, structurations, traitements et usages. Analyse de traces et personnalisation des environnements informatiques pour l'apprentissage humain. In: Hermès, pp. 21–66 (2009)
6. Sébastien, G.: Apprentissage collectif à distance, SPLACH: un environnement informatique support d'une pédagogie de projet. Doctoral Dissertation, Université du Maine (2001)
7. Perry, N.E., Thauberger, C., Hutchinson, L.: gStudy traces of children's self-regulated learning in the Lifecycles Learning Kit. Psychological Test and Assessment Modeling 52, 432–453 (2010)
8. Dyke, G., Lund, K., Girardot, J.-J.: Tatiana: un environnement d'aide à l'analyse de traces d'interactions humaines. Technique et Science Informatiques 29(10), 1179–1205 (2010)
9. May, M., George, S., Prévôt, P.: TrAVis to Enhance Students' Self-monitoring in Online Learning Supported by Computer-Mediated Communication Tools. International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM) 3, 623–634 (2011)
10. Bousbia, N.: Analyse des traces de navigation des apprenants dans un environnement de formation dans une perspective de détection automatique des styles d'apprentissage. Ph.D. Thesis (2011)
11. Djouad, T.: Ingénierie des indicateurs d'activités à partir de traces modélisées pour un Environnement Informatique d'Apprentissage Humain. Ph.D. thesis (2011)
12. Diagne, F.: Instrumentation de la supervision par la réutilisation d'indicateurs: Modèles et Architecture. Ph.D. thesis (2009)
13. Kerkeni, I., Belaid Ajroud, H., Talon, B.: Observation and support in a collaborative pedagogical device: A multi-agent system to control and monitor pedagogical activities. In: Proceeding of the 4th International Conference on Computer Supported Education, Porto, Portugal, vol. 1, pp. 135–140 (2012)
14. Wood, M.F., DeLoach, S.A.: An overview of the multiagent systems engineering methodology. In: Ciancarini, P., Wooldridge, M.J. (eds.) AOSE 2000. LNCS, vol. 1957, pp. 207–221. Springer, Heidelberg (2001)
15. Madkit, `http://www.madkit.org`
16. ILIAS open source e-Learning, `http://www.ilias.de/docu/goto_docu_root_1.html`

MentorChat: A Teacher-Configurable Conversational Agent That Promotes Students' Productive Talk

Stergios Tegos, Stavros Demetriadis, and Anastasios Karakostas

Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece
{stegos,sdemetri,akarakos}@csd.auth.gr

Abstract. MentorChat is a CSCL system that utilizes a conversational pedagogical agent to engage students in fruitful collaboration. This conversational agent attempts to promote students' productive dialogue by providing unsolicited prompts aimed to elicit explicit reasoning. Using MentorChat a teacher is able to design and deploy the phases and the tasks of a collaborative activity along with the support that the agent provides. In this paper, we present (a) the theoretical background MentorChat is based on, (b) the main technical features of the system and (c) some encouraging preliminary results.

Keywords: conversational agents, collaborative learning, productive dialogue.

1 Theoretical Background

Over the past years, researchers have tried to explore the impact of conversational pedagogical agents in both individual and collaborative learning settings [1]. Many conversational agents have been developed to serve as virtual instructors, personal coaches or learning companions, helping students to accomplish several pedagogical goals, such as the acquisition of a foreign language and the development of domain-specific or metacognitive skills [2].

Nevertheless, considering the wide variety of factors a conversational agent system must take into account, these applications are usually considered difficult to implement and often lack the required scalability [3]. Indeed, many systems tend to specialize in specific learning domains, acting as black boxes for teachers or academics, who are expected to use these systems in their curriculum.

Through the prism of Computer-Supported Collaborative Learning (CSCL), our work focuses on the development of domain-independent teacher-configurable conversational agents. From our perspective, the design of an adaptable agent requires the consideration of the following key issues: (a) how we can build a system of low developmental cost that enjoys a broad acceptance by teachers and (b) how we can efficiently trigger students' productive talk by delivering well-targeted agent interventions that increase the probability of constructive peer interactions (e.g. explicit explanations, argumentations) to occur.

Based on the above rationale, we have developed a prototype CSCL system (MentorChat) that enables the teacher to configure dialogue-based collaborative activities, where a conversational agent attempts to scaffold learners' discussion. More specifically,

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 581–584, 2013.

the goal of the agent is to assist learners sustain a 'productive peer dialogue' by encouraging them to explicitly express their reasoning about domain concepts taught in class.

2 The MentorChat System

MentorChat is a responsive cloud-based application requiring a short time to be configured and used in different learning contexts. The system is available for both the English and the Greek language. The web technologies used for its development include MySQL, PHP, AJAX, HTML5 and CSS3. The architecture of MentorChat is displayed in Figure 1.

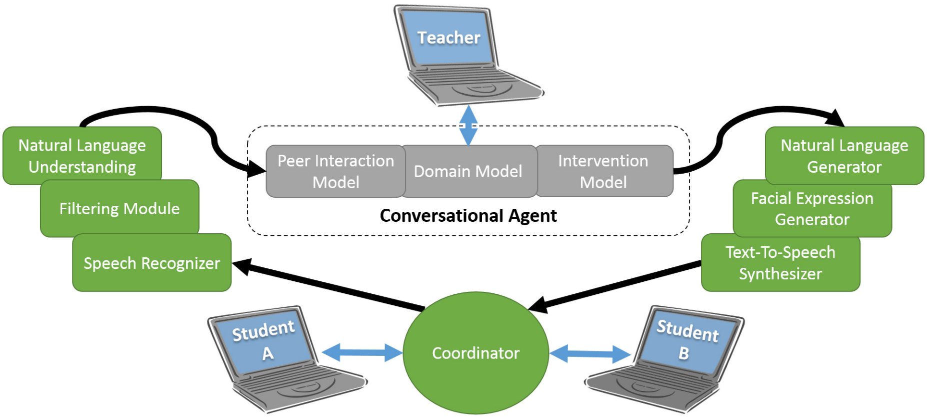


Fig. 1. MentorChat Architecture Diagram

2.1 Student's Interface

The student's interface (Fig. 2) looks like an instant messaging application where students can communicate synchronously either via typing or voice, using the speech recognition feature to compose their messages. When students enter MentorChat, they are asked to collaborate with their partner(s) and provide a joint answer to an open-ended domain question (Fig. 2, A).

During students' discourse the conversational agent may intervene displaying relevant prompts that encourage students to elaborate on key domain concepts. MentorChat interface employs an animated human-like two-dimensional avatar (Fig. 2, B), which serves as the representation of the conversational agent. The agent also utilizes a Text-To-Speech engine for the announcement of its interventions (prompts) in order to enhance the interface modality. Agent interventions are not displayed as common chat contributions (Fig. 2, C) but in a pop-up window (Fig. 2, D), which allows students to have constant access to the agent's last message and respond to it whenever they feel ready.

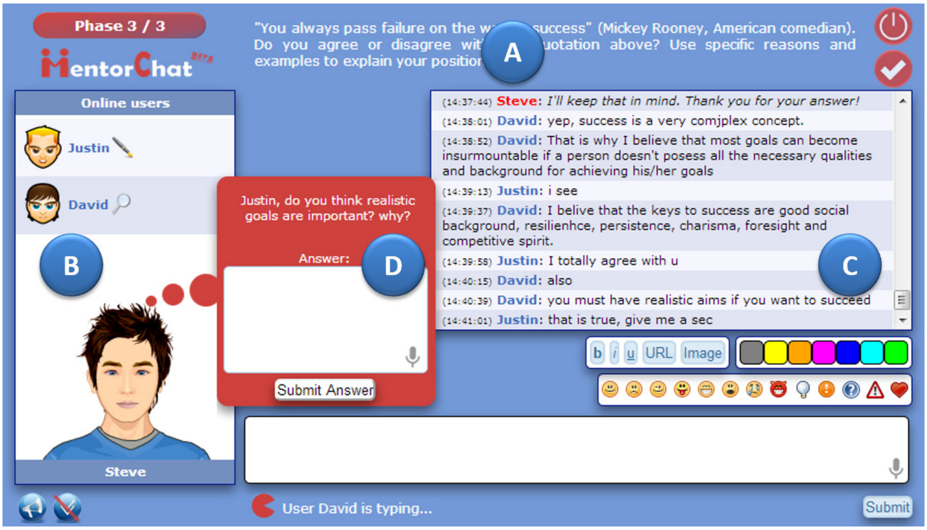


Fig. 2. MentorChat Student's Interface

2.2 Teacher's Interface

A teacher can use MentorChat to set up a dialogue-based activity where students are asked to collaborate in small groups to accomplish a series of tasks. The teacher's interface consists of the following administration panels: (a) the User Management panel, where user, role and group details can be managed, (b) the Activity Structure panel, which enables the teacher to setup the phases/tasks of the collaborative activity, (c) the Monitoring panel, where valuable interaction analysis indicators are displayed at an individual (e.g. number of messages, message frequency) and group level (e.g. on-task rate, number of agent interventions), and (d) the Domain Modeling panel that allows the configuration of the agent knowledge using one of the available modes.

In the basic mode, the teacher is asked to enter only a list of important domain concepts in the form of distinct key-words or -phrases. Based on these concepts, the system is able to synthesize agent interventions using one of the default intervention methods supported by the current version of the system. These include asking students to (a) explicate their reasoning about a specific concept (e.g. "Why do you think that X is important?") or (b) apply their own reasoning to their partner's reasoning (e.g. "Do you agree or disagree with X and why?"). According to the framework of Academically Productive Talk [4], these actions can be regarded as 'productive talk moves', which have proven useful in stimulating constructive classroom interactions.

In the advanced mode, the domain model of the activity is represented by a concept map, which can be formed by entering a set of 'facts'. A fact is a simply structured sentence that consists of a subject, an object and a verb. The system renders these elements dynamically creating two concepts/nodes and their respective relationship/connection in a concept map graph. In the future, we will discuss extensively how the agent can leverage the resulting concept graph to scaffold peers' discourse.

2.3 Conversational Agent

The conversational agent module of MentorChat is based upon three models:

- The peer interaction model, which is responsible for recording and identifying all students' interactions that may serve as opportunities for supportive interventions.
- The domain model, which utilizes the teacher's domain knowledge representation in conjunction with the stemming algorithms and the customizable dictionaries of the system to decide whether an intervention is appropriate.
- The intervention model, which examines a series of various micro-parameters (e.g. the amount of time passed since the last agent intervention, the students targeted by the intervention) to determine if an intervention will eventually be delivered.

3 Results and Outcomes Achieved

Our pilot studies have revealed encouraging evidence regarding the usefulness and usability of MentorChat e.g. [5], [6]. In the context of a university course in multimedia learning, an evaluation study indicated that the agent can improve learning outcomes, helping students to recall and apply the domain terminology more efficiently [5]. Although MentorChat is a work in progress, we consider it to be a valuable tool for experimentation. Its controlled environment and easily tailored actions create new opportunities to explore (a) the best use scenarios that can be suggested to a teacher, who is interested in integrating such a system in real classroom settings, and (b) the impact of various agent triggering policies. For example, our recent study in the Second Language Acquisition domain has shown that an agent intervention strategy targeting a particular student instead of the entire group of students can be more effective in increasing the level of explicit reasoning and productive dialogue [6].

References

1. Gulz, A., Haake, M., Silvervarg, A., Sjödén, B., Veletsianos, G.: Building a social conversational pedagogical agent-design challenges and methodological approaches. In: Perez-Marin, D., Pascual-Nieto, I. (eds.) Conversational Agents and Natural Language Interaction: Techniques and Effective Practices, pp. 128–155. IGI Global (2011)
2. Kerly, A., Ellis, R., Bull, S.: Conversational Agents in E-Learning. In: Proceedings of 28th SGAI International Conference on Artificial Intelligence, pp. 169–182. Springer (2009)
3. Griol, D., Callejas, Z.: An Architecture to Develop Multimodal Educative Applications with Chatbots. International Journal of Advanced Robotic Systems 10(175) (2013)
4. Chapin, S.H., O'Connor, C., Anderson, N.C.: Classroom Discussions: Using Math Talk to Help Students Learn, Grades K-6. Math Solutions Publications (2009)
5. Tegos, S., Demetriadis, S., Karakostas, A.: Exploring the Impact of a Conversational Agent When Triggering Students' Discourse in Online Collaboration. In: Proceedings of 12th International Conference on Advanced Learning Technologies, pp. 321–323. IEEE (2012)
6. Tegos, S., Demetriadis, S., Tsiatsos, T.: A Configurable Conversational Agent to Trigger Students' Productive Dialogue: A Pilot Study in the CALL Domain. International Journal of Artificial Intelligence in Education (under revision)

nQuire for the OpenScience Lab: Supporting Communities of Inquiry Learning

Eloy David Villasclaras-Fernandez, Mike Sharples, Simon Kelley, and Eileen Scanlon

Institute of Learning Technologies, The Open University, UK
Faculty of Science, The Open University, UK
{edvf2,mike.sharples,s.p.kelley,e.scanlon}@open.ac.uk

Abstract. We have developed a platform to support Citizen Inquiry activities, based on the nQuire Toolkit software which was originally designed to support inquiry-based learning activities for schools. Citizen Inquiry is an innovative way for learners to engage in practical scientific activities, in which they take the role of self-regulated scientists in informal learning contexts. nQuire is being integrated with the OpenScience Laboratory, allowing individuals or groups to create inquiries that rely on virtual scientific instruments for collecting scientifically reliable data. A demonstration inquiry has been created using the Open University Virtual Microscope that enables learners to conduct investigations of lunar geology by studying rare and authentic samples of Moon rock collected during the Apollo programme. This demonstration will show the nQuire authoring tools and the prototype inquiry, focusing on the integration of the scientific tool and features that facilitate collaboration in citizen inquiries.

Keywords: Citizen inquiry, inquiry based learning, nQuire toolkit, OpenScience Laboratory, Geology of the Moon.

1 Background (Pedagogy)

The goal of the TEL system presented in this demonstration is to support an innovative pedagogical approach to inquiry science learning: **citizen inquiry** [2]. Citizen inquiry refers to the design and enactment of scientific projects by non-professional scientists within a community of users. From a pedagogical point of view, citizen inquiry follows a similar approach as inquiry based learning [1]; however, citizen inquiry is driven by personal interest and is not related to a course curriculum.

At the same time, citizen inquiry integrates elements of open science [4] and citizen science [3]. Citizen inquiry shares with the latter the emphasis on involving members of the public who may have not received training as researchers. However, citizen inquiry aims at developing their interest in science by making them responsible for the planning, management and realization of complete scientific projects.

Supporting citizen inquiry poses a number of interesting research challenges:

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 585–588, 2013.

- **Challenge 1: Learners' Motivation.** Citizen inquiry does not have the inherent motivation of inquiry based learning activities, nor that of citizen science projects. Our goal is to generate motivation through the creation of a user community that includes people interested in science with widely varying levels of expertise. Interaction with other users, the oportunity of helping or receiving help from experts, collaborative inquiries, and recognition within the citizein inquiry community are some of the features that are expected to motivate participants to join the community.
- **Challenge 2: Adopting a Process of Inquiry.** Supporting users through inquiry processes is a relevant problem of inquiry based learning. In citizen inquiry, in which the figure of a teacher or tutor is absent, the same issue may be problematic. This is related to one of the learning goals behind citizen inquiry: to facilitate learners acquire an understanding of scientific methodologies.
- **Challenge 3: Supporting Scientifically Relevant Activities.** While our goal with respect to citizen inquiry is to facilitate learning, we are interested in enabling non-professionals to carry out scientifically relevant and accurate research, that adheres to the methodological and ethical principles of the scientific activity.

These challenges have been tackled by the development of a TEL system. To demonstrate these features, a first prototype of a citizen inquiry has been created: the *Moon Rock Explorer*. It is an open, self-regulated investigation of lunar geology, designed for learners without large experience in Geology.

2 Background (Technology)

To support citizen inquiry activities, several extensions to the nQuire Toolkit have been made:

- **Integration of Scientific Instruments.** nQuire is currently being integrated with the OpenScience Laboratory (OSL). The OSL will provide access to virtual and remote scientific tools, such as virtual microscopes, Treezilla, the PIRATE remote telescope, and physics equipment to conduct experiments. These intruments allow the collection of valid scientific data, which can be used in the frame of scientific research. Additionally, all the instruments can be accessed using a web browser. Application Programming Interfaces are being developed and tested to enable the creation of inquiries through the nQuire Toolkit that make use of these instruments to collect data. Figure 1 shows the Virtual Microscope data collection interface within nQuire, which is used in the *Moon Rock Explorer* inquiry to investigate four samples of Moon rock collected by Apollo astronauts.
- **Support for Collaborative Self-regulated Inquiry.** The *Moon Rock Explorer* is an individual activity. Nevertheless, it includes support for interaction between users: they can publish and share their investigations through automatically created forum topics. This feature is designed to facilitate the discussion of investigations, enabling experts to support participants, etc.

– **Support for Self-management of Inquiries.** The current version of nQuire allows any user registered in the system (not only teachers) to create their own inquiries, with any of the available scientific instruments. They can decide to share them with friends, or with anyone in the system. Each participant in an inquiry can be assigned to roles (e.g., *Project leader*, *Data collector*, etc.) to control the access to each inquiry activity. The authoring tool is designed to support the creation of inquiries with features of inquiry based learning activities, citizen science projects, and citizen inquiry.

The integration of nQuire with the OpenScience laboratory requires the definition of a series of Application Programming Interfaces (APIs) that enable communication between scientific instruments and nQuire. Currently, as shown in this demonstration, only the Open University Virtual Microscope has been linked to nQuire.

3 Results and Outcomes Achieved

The first prototype of the *Moon Rock Explorer* inquiry has been evaluated with PhD students in an informal learning context. The evaluation is described in [2]. The goals of the study were to assess the relevance of the challenges discussed in Section 1, and the technological support described in Section 2.

The evaluation showed that the nQuire Toolkit provides adequate support for citizen inquiry, even when the figure of the teacher is no present. Similarly, the results indicate that the integration of the Virtual Microscope, shown in Figure 1, was also successful. The feedback provided by the participants allowed us to improve the virtual microscope data collection interface, and points us to new features to support motivation and user interaction.

The evaluation has thus served to validate two parts of the system: inquiry process guidance, and access to the Virtual Microscope. Based on these results, we will proceed to integrate further scientific instruments. The evaluation has also provided insights as to the requirements to support community building around scientific inquiry, which are leading to further improvements of the system.

4 Demonstration Outline

The demonstration shows the complete nQuire Toolkit system integrated with the OpenScience Laboratory. The first part of the demonstration will focus on the authoring tool, including: (1) Structuring citizen inquiry by creating phases and activities; (2) Creating data collection activities that integrate external scientific instruments; and (3) Specification of collaboration modes, including management of groups and participants roles within the inquiry.

The second part will deal with the runtime system. To illustrate it, the *Moon Rock Explorer* inquiry (see Section 3) will be presented. The demonstration will show briefly the whole process that the learners may go through, including the investigation of Moon rock samples using the Virtual Microscope.

The guests will be able to use the system, taking two different roles: inquiry designer, and inquiry participants.

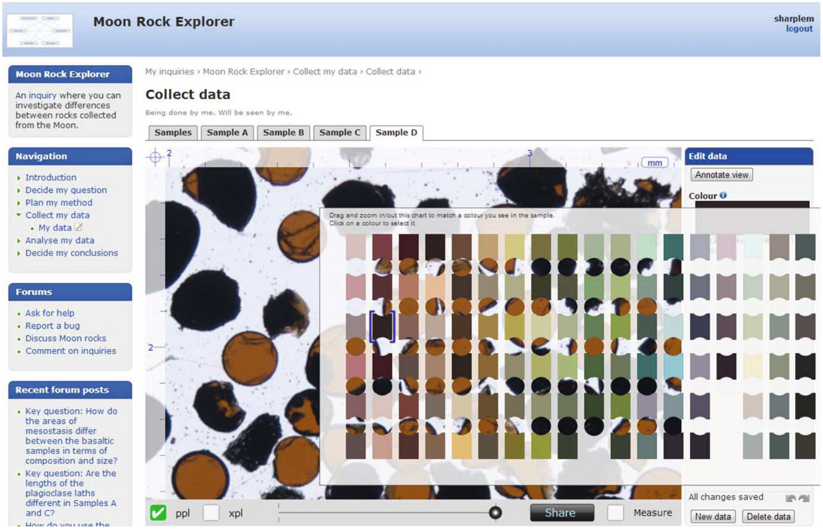


Fig. 1. Virtual Microscope data collection interface in nQuire

Acknowledgements. This work is developed within The OpenScience Laboratory project, funded by The Open University and the Wolfson Foundation. The authors would like to thank the members of OU Institute of Learning Technology that participated in the evaluation of the Moon Rock Explorer inquiry; and members of the Personal Inquiry project from University of Nottingham and the Open University for their work the nQuire.

References

1. Dewey, J.: Science as Subject-matter and as Method. Science 31(787), 121–127 (1910)
2. Villasclaras-Fernandez, E., Sharples, M., Kelley, S., Scanlon, E.: Supporting Citizen Inquiry: An investigation of Moon Rock. In: Hernández-Leo, D., Ley, T., Klamma, R., Papadopoulos, G. (eds.) EC-TEL 2013. LNCS, vol. 8095, Springer, Heidelberg (2013)
3. Wilderman, C.: Models of community science: Design lessons from the field. In: McEver, C., Bonney, R., Dickinson, J., Kelling, S., Rosenberg, K. (eds.) Citizen Science Toolkit Conference. Cornell Laboratory of Ornithology, Ithaca (2007)
4. Woelfle, M., Olliaro, P., Todd, M.H.: Open science is a research accelerator. Nature Chemistry 3, 745–748 (2011)

On Self-adapting Recommendations of Curricula for an Individual Learning Experience

Sebastian Bab and Luise Kranich

Forschungszentrum Informatik (FZI), Germany
{bab,kranich}@fzi.de

Abstract. We propose ideas for the development of TEL systems which allow for an automatic, dynamic and self-adapting recommendation of curricula from a wide set of available content for an individual user and with regard to a specific purpose. We argue that recommender systems in the prevalent occurrence cannot be used directly in TEL systems, but must be extended by process-related techniques for continuous optimization and adaptation of the generated curriculum.

1 Making TEL Personal – Problem Statement

There are currently no implementations of recommender systems in TEL that actually generate dynamic learning paths leading to a specific goal while focusing on the individual learner's needs. The use of collaborative filtering, item-based filtering or a hybrid of both will not be satisfactory in a learning environment where units should be consumed in a certain order, and where learning paths aim towards reaching a certain educational goal [2]. The best path towards this goal is based on a learner's individual learning style and a didactic concept that will challenge but not overburden users.

2 Process-Based Filtering – Approach to Solution

Sequence and orientation are two of the main characteristics that distinguish recommender systems in TEL environments from those used by online shops and music or video platforms: Not the content consumption itself serves as a goal to be maximized, but a meaningful combination of items forming a path towards a pre-defined purpose. The central idea of our recommender system with process-based filtering is given by the combination of dynamic user profiles with pre- and postconditions of available content items regarding certain learning purposes: Each content item in the TEL system represents a learning unit in a specific field with a specific level of detail. A precondition of a content item defines the minimal set of competences needed in order to successfully complete the unit. A similar concept can also be found in Competence-Based Knowledge Space Theory [1]. A precondition defines for each competence the degree of expertise needed in this competence for successfully passing the content item (compare

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 589–590, 2013.

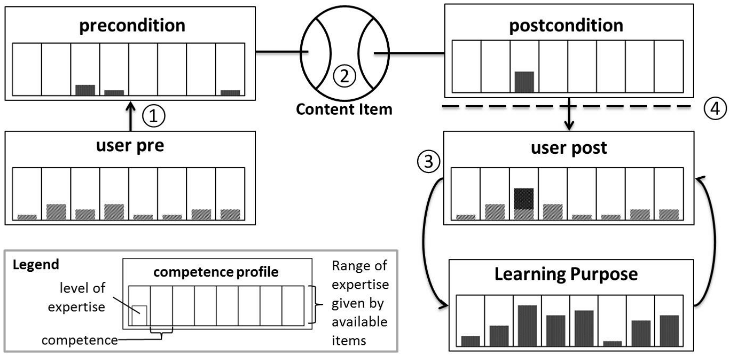


Fig. 1. Content items are sorted by defining pre- and postconditions matched with the learner's competence profile (1). The recommended item is chosen out of the group of items that lead to the learning purpose (2+3). To determine how the consumption of a content item has added in the intended direction, an evaluation is implemented (4).

with Figure 1). The postcondition describes the competences maximally achievable by successfully completing the unit. The depth of understanding can be determined by formalized tests or exams or less explicit methods.

A state of a dynamic learner profile is given as a set of numeric values describing the learner's current expertises in the regarded competences. The learning purpose is defined in the same way and thus represents a virtual user profile. Based on pre- and postconditions, the system will identify the content equivalence classes which the user is capable of passing and thereby directing him towards the learning purpose. Which content item is chosen from the generated set can be determined by classic filtering methods of recommender systems. Regarding the overall learning process, didactical methodologies can be implemented by additional filters further narrowing the set of recommendation candidates.

3 Conclusion

We argued that our described process-based filtering can be a basis for implementing a TEL platform that is capable of recommending adaptive dynamic learning paths which can lead to a much more personalized learning experience.

References

1. Albert, D., Hockemeyer, C., Kickmeier-Rust, M.D., Nussbaumer, A., Steiner, C.M.: E-learning based on metadata, ontologies and competence-based knowledge space theory. In: Lukose, D., Ahmad, A.R., Suliman, A. (eds.) KTW 2011. CCIS, vol. 295, pp. 24–36. Springer, Heidelberg (2012)
2. Ley, T., Kump, B., Gerdenitsch, C.: Scaffolding self-directed learning with personalized learning goal recommendations. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 75–86. Springer, Heidelberg (2010)

A Generative Computer Language to Customize Online Learning Assessments

Antonio Balderas, Iván Ruiz-Rube, Juan Manuel Dodero, Manuel Palomo-Duarte, and Anke Berns

University of Cádiz, Spain, Cádiz, C/Chile 1, CP 11002
{antonio.balderas,ivan.ruiz,juanma.dodero,manuel.palomo,anke.berns}@uca.es

Abstract. The focus on assessment of learning experiences has shifted from knowledge to competences. Unfortunately, assessing certain competences is mainly a subjective task, being problematic for both the evaluators and the evaluated. Additionally, when the learning process is computer-supported and the number of students increases, traditional assessment procedures suffer from scalability problems. In this paper we propose a query language that supports grading learning competences according to students' performance in an online course. Using it we automatically extract different objective indicators about students work in a Learning Management System (LMS). Evaluators can use this computer programming-like language to express a number of required indicators. Such indicators are automatically obtained from the activity logs generated by the LMS.

Keywords: online learning, competence assessment, technological support in online education, domain-specific languages, learning analytics.

1 Introduction

The scope of our work is related to the virtual places specially designed for collaborative tasks on LMS, which they are widely used nowadays. Each file, access or assignment done by a student is registered in the system [1]. This collected information, properly filtered [2], would be used as indicators of students' work.

2 EvalCourse

EvalCourse is the domain-specific language [4] that we have developed to get indicators. Language syntax and reserved words are listed in 1.1. The first line specifies the name of the evidence or indicator. The second line is a constant that indicates that we want to get information about students. The third one indicates the kind of information to be retrieved: milestones, participation or access. And the last one if it is collected from assignments, forums or campus activity. In case one or more activity identifiers are provided ([list of ids]), only information on those assignment, forum or campus will be provided. If not specified, the query acts on all activities of the indicated type.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 591–592, 2013.

```
Evidence name_of_the_indicator:
        get students
        show milestones | participation | access
        in assignment | forum | campus [list of ids].
```

Listing 1.1. Language syntax. Reserved words are highlighted in bold format.

We will illustrate language usage with a simple example on assessing problem solving competence. We can assess students' performance in this competence using their participation in forums [3] in which lecturers set out problems. We ask for the information with the code shown in 1.2, obtaining results in an XML file that can be imported in most popular spreadsheets.

```
Evidence Participation_list:
        get students
        show participation in forum.
```

Listing 1.2. Code to retrieve students' performance in forum participation

3 Conclusions and Future Work

Using our system, the lecturer, without any technical knowledge on system programming and using a very simple syntax, can automatically assess the performance in several competences for any student, justifying it with objective figures. At the moment, our system only works with an instance of Moodle, so our future work is implement support for other LMS.

Acknowledgements. This work has been funded by the *PAIDI* programme of the Andalusian Government, ASCETA project (P09-TIC-5230) and University of Cadiz programme for Researching and Innovation in Education.

References

1. Chebil, H., Girardot, J., Courtin, C.: An ontology-based approach for sharing and analyzing learning trace corpora. In: Proceedings - IEEE 6th International Conference on Semantic Computing, ICSC 2012, pp. 101–108 (2012)
2. Florian, B., Glahn, C., Drachsler, H., Specht, M., Fabregat Gesa, R.: Activity-based learner-models for learner monitoring and recommendations in moodle. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 111–124. Springer, Heidelberg (2011)
3. Jacob, S.M., Lee, B., Lueckenhausen, G.R.: Measuring critical thinking skills in engineering mathematics using online forums. In: 2009 International Conference on Engineering Education, ICEED 2009 - Embracing New Challenges in Engineering Education, pp. 225–229 (2009)
4. van Deursen, A., Klint, P., Visser, J.: Domain-specific languages: an annotated bibliography. SIGPLAN Not. 35(6), 26–36 (2000)

Monitoring Learning in Children with Autism

Valentina Bartalesi[1], Maria Claudia Buzzi[2], Marina Buzzi[2],
Barbara Leporini[1], and Caterina Senette[2]

[1] ISTI-CNR, via Moruzzi 1, 56124 Pisa, Italy
{Valentina.Bartalesi,Barbara.Leporini}@isti.cnr.it
[2] IIT-CNR, via Moruzzi 1, 56124 Pisa, Italy
{Claudia.Buzzi,Marina.Buzzi,Caterina.Senette}@iit.cnr.it

Abstract. We present a learning analytic tool for analyzing data collected during didactic sessions performed by children with autism, via software. The tool is a web application that automatically extracts, aggregates and visualizes the children's performance data collected when they perform technology-enhanced exercises during didactic activities. The aim is to provide teachers with easy real-time monitoring of learning progress or difficulties over time, thus enabling continuous and personalized tuning of the didactic intervention. Extracted data can be visualized through tables and graphics, allowing the user to further explore them in an interactive modality. Educators were involved in the design phase of the tool, in order to define the most important analysis parameters to extract from the huge amount of collected data. The educators were also involved in the test phases to improve the system's usability.

Keywords: autism, learning, ABA, data analysis, learning analytic tool.

1 Introduction

Autism Spectrum Disorder (ASD) is classified as a type of Pervasive Developmental Disorder that affects individuals with different degrees of impairment. Traditional education methods are ineffective for children with autism, so different teaching approaches aimed at better exploiting the subject's abilities are adopted. Of these, Applied Behavioral Analysis (ABA) has proven effectiveness [1]. It provides one-to-one teaching via trials and exercises of increasing difficulty. The child's progress is constantly monitored by collecting data from the child's performance of the trials, which then need to be analyzed.

2 The Learning Analytic Tool

ABCD SW is an open source web application that supports learning for children with autism -- age range 2-6 years – by implementing basic ABA programs [2]. The ABCD SW modules allow the child to learn and become familiar with categories such as 'food', 'colors', 'animals', 'letters', and their articles such as 'pear', 'blue', 'cat', 'd'. The software automates the exercises while it enables recording the children's

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 593–594, 2013.

performance data. In order to provide teachers with an easy real-time tool for monitoring the children's learning progress/difficulties over time, we have developed a learning analytic tool to process the large amount of data collected using ABCD SW. The tool, a web application implemented in PHP, JavaScript and jQuery, involved teachers in its design to better monitor the children's learning, thus enabling continuous and personalized tuning of the didactic intervention. The tool offers a basic set of queries that can be arduous to perform manually. The query results are presented in table and chart format: stacked bar, multi-line and single-line. We used Highcharts, a library written in JavaScript, which offers intuitive and interactive chart for web applications. By default, a chart can be exported and downloaded as a PDF file, as an image, in PNG or JPEG formats, or as an SVG representation. It is also possible to export and download row data into CSV format to refine the data view as preferred. The tool allows the tutor to choose from among six different predefined queries: (i) the list of all the articles and (ii) categories utilized in the didactic sessions; (iii) number of articles mastered by the child; (iv) the percentage of correct trials performed by the child; (v) the percentage of different kinds of error; (vi) the percentage of occurrence of different kinds of prompts. For each query the user can select a child and set a time interval.

Fig. 1. a) Graph of number of articles mastered vs time **b)** Graph of error types vs time

In Fig. 1 two examples for the queries (iii) and (v) are shown in a graphic format. The data analysis tool is open source and freely available. Suggestions and comments collected from 15 users showed that the tool is greatly appreciated by ABA professionals since it eliminates the effort of manually retrieving the large amount of data produced in ABA sessions in order to evaluate the child's performance.

References

1. Rogers, S., Vismara, L.: Evidence-based comprehensive treatments for early autism. Journal of Clinical Child & Adolescent Psychology 37, 8–38 (2008)
2. Buzzi, M.C., Buzzi, M., Gazzé, D., Senette, C., Tesconi, M.: ABCD SW: Autistic behavior & computer-based didactic software. In: Proceedings of the Int. Conference on Web Accessibility (W4A 2012), Article 28 , 2 pages. ACM, New York (2012)

From Geocaching to Mobile Persuasive Learning – Motivating the Interest in the Life and Work of Danish Author Kaj Munk

Lasse Burri Gram-Hansen, Sandra Burri Gram-Hansen, and Peter Øhrstrøm

Royal School of Library and Information Science, University of Copenhagen,
and Department of Communication and Psychology, Aalborg University
lgh@iva.ku.dk, {burri,poe}@hum.aau.dk

Abstract. This paper presents some of the initial steps taken towards digital mediation of the cultural heritage related to Danish author Kaj Munk and the impact these steps have had on the ongoing research on persuasive learning.

Keywords: Persuasive Learning, Cultural heritage, Mobile technology.

1 From Geocaching to Mobile Persuasive Learning

This paper presents and discusses how promising results from a location-based experience design in a Danish museum are being taken into consideration in the development of a *persuasive learning* design which aims to motivate and engage learners through a constructivist approach to learning, and the use of multimodal learning technologies.

The novel concept of persuasive learning has been explored and developed in the past few years, by the EuroPLOT[1] research project [1]. EuroPLOT is funded by the Life Long Learning programme, and seeks to investigate the theoretical and practical cross-field between learning technologies and persuasive design, as initially introduced by Fogg [2]. The concept of persuasive learning is being tried within four highly diversive work cases all sharing the distinction that they each focus on highly complex learning material and include topics such as language learning, chemical substance handling, academic business computing and archival based studies.

The latter of these cases focuses on motivating learners to develop a deeper understanding of the life and works of Danish vicar, author and playwright Kaj Munk who was assassinated by Gestapo in January 1944 after having aired his resistance against the German occupants through his sermons and his literary works. As a result, Kaj Munk has become an important figure not only in Danish literature, but also in Danish history, and the mediation of his life and works receives attention both in learning and in mediation of cultural heritage.

A 2006 pilot study from Thorning elementary school, Denmark, used mobile tech- nology to enable teaching outside the classroom [3]. Using the location-based

[1] http://www.eplot.eu/

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 595–596, 2013.

GPS- hunt Geocaching[2], school teachers encouraged students to learn different subjects in appropriate surroundings, leveraging the students' fascination with technology to spark their interest. Inspired by this study as well as research investigating the potential of Geocaching in a persuasive perspective [4], a system of four geocaches related to Kaj Munk has been established in cooperation with the Kaj Munk Museum in Vedersø, which is located in Kaj Munks former vicarage. When locating a geocache, users provide digital feedback, which has provided important insight towards the impact of geocaches in relation to motivating a further interest in Kaj Munk. At present, the geocaches in Vedersø have generated more than 400 logs.

Inspired by the use of Geocaching in learning scenarios as well as other similar experiments and experiences, it has been discussed in EuroPLOT how mobile persuasion can be used in relation to a teaching task like the one related to Kaj Munk's life and works. It has been argued that mobile devices hold potential which may ease the task of incorporating the surroundings in a learning design, and as a result one of the learning technologies developed within the project, now provides teachers with the ability to create not only traditional but also mobile learning objects which can be executed through smartphones and tablet computers.

2 Mobile Persuasive Learning Put to Practice

In order to test both types of learning objects, a persuasive learning design has been developed for the students of the lower secondary school in Vester Hassing. This design included both traditional and mobile learning objects all of which were linked to an overall intended learning outcome. The use of mobile learning objects is particular relevant to the Kaj Munk case due to the connection between historical events, literature and specific locations, which constitute a particular distinction for this work case. Results of the pilot study were promising. Both in terms of the students meeting the intended learning outcomes, and in terms of becoming engaged in the learning process and increasingly motivated to learn more about the subject.

References

1. Gram-Hansen, S.B.: Persuasive Learning Design Framework. EuroPLOT Deliverable, vol. 3.3. Aalborg Universitet, Aalborg (2012)
2. Fogg, B.J.: Persuasive Technology - Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco (2003)
3. Thomsen, T.B. GPS i Undervisningen (2008)
4. Gram-Hansen, L.B.: Geocaching in a persuasive perspective. In: Proceedings of Poster Papers for the 4th International Conference on Persuasive Technology. ACM, Claremont (2009)

[2] http://www.geocaching.com/

Detecting Discourse Creativity in Chat Conversations

Costin-Gabriel Chiru and Traian Rebedea

University Politehnica of Bucharest, Department of Computer Science and Engineering, 313 Splaiul Independetei, Bucharest, Romania
{costin.chiru,traian.rebedea}@cs.pub.ro

Abstract. In this paper we propose a new method for identifying creativity that is based on analyzing a corpus of chat conversations on the same topic and extracting the new ideas expressed by participants. The application is a first step in supporting creativity in online group discussions by highlighting the novel concepts present in conversations (*new ideas*) and also by identifying topics that could have become important, if not forgotten during the debates (*lost ideas*).

Keywords: Creativity, Linguistics, CSCL, Learning Analytics.

1 Introduction

Creativity is defined as the ability to transcend traditional ideas, patterns, relationships into meaningful new ideas, interpretations, etc. [4]. In written discourse, it is called linguistic creativity [3] and it measures "new and creative ways of expressing a given idea". Linguistic creativity is sometimes described as an unpredictable departure from the rules of regular word formation [2]. In our opinion, linguistic creativity could be defined as a deviation from the standard way of writing on different topics.

We propose a new approach for assessing discourse creativity based on processing a large volume of chat conversations on the same topic. Thus, we see the rare concepts (that are not off-topic) as being intentional deviations from the standard concepts that are debated in the corpus of chats, representing a sign of discourse and linguistic creativity. This way we link creativity to *differences in the participants' discourses.*

2 Corpora and Application

The corpora used for analysis consists of chats of students that worked in small teams to find the best online instrument for information sharing in a company [1]. We analyzed the content of 62 chat conversations in order to detect significant similarities and differences between them. We considered that the similarities are represented by common concepts that were uttered by the majority of the participants, while the differences – concepts present only in very few conversations – should be signs either of creativity or divergent thinking or of off-topic content. In order to discriminate between these two opposite situations, we evaluated the utterances to detect which ones were on topic. Therefore, we considered that the on-topic words were the ones from the lexical field of concepts such as "blog", "chat", "forum", "wiki", or "web".

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 597–598, 2013.

2.1 Lost Ideas, Similar Ideas, Ideas and Reactions

The *lost ideas* represent the concepts found only in a small number of chats. For their detection, we first identified the rare words from the corpus (the ones found only once) and afterwards checked to see if they were on-topic by evaluating the utterances containing them (whether they also contained on-topic concepts or not). *New ideas/similar ideas* express concepts that have a high frequency in a small number of chats. For each utterance where we found such a concept, we extracted a text pattern (a window of type "$word_1$ *[*]* $word_2$") and used this pattern to detect similar situations in the other chats.

We have also devised a method that classifies the participants to chats into 5 classes according to their discourse creativity: leaders, developers, innovators, negativists, and conclusives. This list is open-ended so there may also be other typologies. In order to detect the typology of a participant, it is first necessary to automatically classify the chat utterances according to their content in the following 5 different classes:

- *Ignored utterances* – usually they contain social noise.
- *Ideas* – utterances containing concepts that are on-topic.
- *Conclusions* – contains patterns that are specific to conclusive utterances.
- *Developments to conclusions* – if the previous utterance contains some on-topic concepts, its development should contain words from the same specter.
- *Reactions* – approvals, disapprovals or continuations of previously stated ideas.

3 Conclusions

Discourse creativity is a concept that is very difficult to assess automatically. In this paper, we have presented a method for determining creativity in a large collection of online discussions that debate the same topic within a CSCL scenario. Moreover, the proposed method can be used to identify typologies of participants starting from the elements used for assessing creativity.

References

1. Rebedea, T., Dascalu, M., Trausan-Matu, S., Armitt, G., Chiru, C.: Automatic Assessment of Collaborative Chat Conversations with PolyCAFe. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 299–312. Springer, Heidelberg (2011)
2. Renouf, A.: Tracing lexical productivity and creativity in the british media: the chavs and the chav-nots. In: Lexical Creativity, Texts and Contexts, pp. 61–89. John Benjamins Publishing Company, Amsterdam (2007)
3. Veale, T.: Creative language retrieval: A robust hybrid of information retrieval and linguistic creativity. In: Proceedings of ACL 2011, Portland, Oregon, USA, pp. 278–287 (2011)
4. Zhu, X., Xu, Z., Khot, T.: How creative is your writing? a linguistic creativity measure from computer science and cognitive psychology perspectives. In: Proceedings of the Workshop on Computational Approaches to Linguistic Creativity, pp. 87–93. ACL (2009)

IEEE-IST Academic: A Web-Based Educational Resources Case-Study

Ana Moura Santos[1] and Rui Costa[2]

[1] Department of Mathematics, Instituto Superior Técnico,
Technical University of Lisbon,
Av. Prof. Doutor Anibal Cavaco Silva, 2744-016 Porto Salvo
amoura@math.ist.utl.pt

[2] IEEE-IST Student Branch, IEEE Portugal Section
ruicosta@ieee.org

Abstract. The project IEEE-IST Academic is an initiative of the IEEE Student Branch at Instituto Superior Técnico (IST) and since its beginnings (May 2012) counts with the collaboration of several professors of IST. Organized by course topics, the first Academic videos offer all undergraduate students at IST complementary materials to most of the Mathematics undergraduate courses, Computer Sciences Introductory Course, Chemistry and Electronics. The students technical team is responsible for the capture, processing and edition of the materials, and is also in charge of the design and back office of the IEEE-IST Academic portal. The videos are designed to be accessed from several platforms, such as: laptops, tablets, smartphones, and extensively make use of visualization and 3D graphics 3D. The Academic videos are designed to be more close to the courses syllabus at IST, but can easily be a future reference for other Portuguese, Brazilian and European schools of Engineering, Science and Technology.

Keywords: web-based education, online contents, teacher-student collaboration, educational web-portal.

1 The IEEE-IST Academic Project – A Description

On October 2011 the IEEE-IST Academic project started being drafted, coordinated by one of the authors, Rui Costa, Master-degree student at IST and Educational Activities Chair at the IEEE Portugal Section. Contacts were made with Nape-TP (Students Support Office at IST-Taguspark) for support in connection with students grants for the team and project promotion among the teachers. Ana Moura Santos, the other , being head of the Nape-TP and a math teacher, facilitated the contacts with fellow professors at the same institution. The decision was made that top priority were the topics for first years undergraduate students at IST. On the 25th of May, the IEEE-IST Academic website[1] was first made available to the public counting with 69 videos, having today more than 200 of several

[1] http://academic.ieee-ist.org/

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 599–600, 2013.

different topics. The student project team designed a website having in mind principles and requirements such as, clean, adaptive and simple interface, quick access to information of a specific course and ability to search and find relations between contents, to be used in both computer and mobile devices.

Being an initiative designed and developed by students, the IEEE-IST Academic is aimed to closely cope with the students practices of networking [2] and academic student needs at IST. The result is a choice of video formats that maximize the video effectiveness of the different available technologies: **traditional**, video modules which are recorded in front of a board, with a teacher presenting a given topic; **tablet**, contents that are explained by using a tablet to write down notes on top of pictures, schemes or diagrams; **on-field**, video modules that are recorded on specific locations, e.g. demonstrating a chemical reaction in a lab; **screencast**, designed for demonstrations made in a computer, where all the screen is captured while the teacher explains focusing on the relevant parts.

The Academic videos are not specialized one-off courses offered by professors, that substitute any course, in contrast with the Massive Open Online Courses (MOOCs)[1][2]. Instead they constitute complementary topics, sometimes with basic concepts and applications, sometimes with carefully chosen proofs and sophisticated algorithms, mostly planned having in mind difficult or crucial moments in the learning process. The granular structure of the videos is similar to the videos of Khan Academy[3], the choice of the native language: Portuguese, since despite the fact that several widespread initiatives offer web accessible academic contents[4], some students find hard to use those contents in a non-native language, with no direct translation to the technical local jargon.

According to a online survey conducted after the exams period of June 2012, students say that the Academic contents help them to prepare for laboratories, homework's, quizzes and evaluation periods. Taking into account the period between September 2012 and February 2013, equivalent to one full semester of classes in Portugal, the website had a total of 16 959 visits from 5 871 unique visitors, with a total of 58 041 pageviews, meaning a rate of 3,42 pages views per visit. The average visit duration was of 5 minutes and 48 seconds. During this period, from all the visits registered, 65.53% were returning visitors while 34.47% were new visitors.

References

1. Daniel, J.: Making sense of MOOCs, Academic Partnerships (2012)
2. Rennie, F., Morrison, T.: e-Learning and Social Networking Handbook, 2nd edn. Routledge, N.Y (2012)
3. Khan, S.: Let's use video to reinvent education, TEDGlobal (March 2011), `http://www.ted.com/talks/`
4. Koller, D.: What we're learning from online education, TEDGlobal (June 2012), `http://www.ted.com/talks/`

[2] The Bill and Melinda Gates Foundation recently announced it is funding a MOOC Research Hub, what reveals that peer-reviewed research on MOOCs has been minimal.

Semantic Social Sensing for Improving Simulation Environments for Learning

Vania Dimitrova[1], Christina M. Steiner[2], Dimoklis Despotakis[1], Paul Brna[1], Antonio Ascolese[3], Lucia Pannese[3], and Dietrich Albert[2]

[1] School of Computing, University of Leeds, United Kingdom
[2] Knowledge Technologies Institute, Graz University of Technology, Austria
[3] Imaginary Srl, Milan, Italy

1 Overview of Semantic Social Sensing

The rapidly growing learning simulations market calls urgently for innovative ways to facilitate the simulation design process [1],[2]. Social spaces can provide an extensive source of reports on individuals' experiences and their real-world contexts that may be exploited for the purpose of identifying relevant content and evaluating the quality of a simulation. To realise this potential, appropriate ways to make sense of user generated content (UGC) are needed. This work presents a novel approach, called *semantic social sensing* (SSS), which exploits ontologies and semantic augmentation combined with discourse analysis uncovering intentions behind the user comments. We have developed two SSS instruments enabling analysis of UGC – (a) a framework for automatic semantic analysis for capturing viewpoints (ViewS), which utilises ontologies and semantic tagging and enrichment and enables visual exploration of the conceptual spaces associated with UGC [3]; and (b) a schema for discourse analysis to identify intentions useful for simulator design [2] and inspired by research analysing communicative functions of user contributions in collaborative settings [4].

2 Empirical Application: Simulator and User Trial

A user trial was conducted from October 2012 until January 2013 involving 39 participants using a simulator for promoting intercultural awareness in interpersonal communication (IC) in business settings, developed by imaginary Srl within the framework of the ImREAL[1] EU project. The simulated scenario covers four episodes each including several situations: *Greeting*, *Dinner*, *Bill*, and *Goodbye*. The simulator is equipped with a microblogging tool, which enables UGC to be collected *in-situ*.

The 206 microposts collected in the study were semantically augmented and analysed with ViewS, combined with discourse analysis carried out by three independent annotators. *Agreed categorizations* for comments were derived following the majority principle. The SSS analysis output is outlined in Table 1. The semantic maps produced with ViewS indicated that the *simulation experience had triggered user comments linking to a range of social signals* and UCG could be grouped in several meaningful clusters. Discourse categories provided another way for grouping

[1] `http://www.imreal-project.eu`

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 601–602, 2013.

of comments, which were then linked to semantic analysis with ViewS for further zooming into the discourse to get an insight into the related UGC.

Table 1. Results based on the ViewS semantic annotation and discourse analysis

Instrument		Greetings	Dinner	Bill	Goodbye	Total
ViewS	**WNAffect**	36	36	11	5	57
	Body Language	76	63	43	33	106
Discourse schema	**Real world stories**	16	13	6	1	36
	Rules	54	38	11	7	110
	Idiosyncrasies, stereotypes and beliefs	8	22	3	4	37
	Affective judgement related to the simulated situation	7	8	3	2	20
	Statements about the situation	18	30	10	6	64

Overall, the application of SSS in the specific use case gives an initial indication of the suitability and usefulness of the approach and its instruments to address requirements and needs for designing learning simulators. Most added value appears to be on organizing, grouping and comparing UGC presenting real world stories, personal experiences, and rules. This can be helpful for extending the simulation content (e.g. new situations, branches, additional descriptions), improving the feedback (e.g. including more options, referring to shared experiences), extending the options given to the learner. Furthermore, in some cases, SSS output can be used as reassurance that relevant concepts have been recognized or recalled, or as confirmation that simulated situations relate to real world experiences.

Future work will aim at improving and validating the two instruments. To further support simulation designers, we plan to increase the level of interaction between users and ViewS and to enable seamless integration into in-house tools for simulator design. The discourse schema requires further extension and elaboration to increase reliability and wider applicability to experiential learning environments.

Acknowledgements. The research leading to these results has received funding from the EU 7th Framework Programme under grant ICT 257831 (ImREAL project). We would like to thank the users taking part in the trial and the three annotators.

References

1. Swartout, W.: Lessons learned from virtual humans. AI Magazine 31, 9–20 (2010)
2. Hetzner, S., Steiner, C., Dimitrova, V., Brna, P., Conlan, O.: Adult self-regulated learning through linking experience in simulated and real world: A holistic approach. In: Proc. of EC-TEL 2011, pp. 166–180 (2011)
3. Despotakis, D., Dimitrova, V., Lau, L., Thakker, D., Ascolese, A., Pannese, L.: ViewS in user generated content for enriching learning environments: A semantic sensing approach. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS, vol. 7926, pp. 121–130. Springer, Heidelberg (2013)
4. Arvaja, M., Pöysä-Tarhonen, J.: Tracing discursive processes of shared knowledge construction in a technology-enhanced higher education setting. Interactive Learning Environments, 1–17 (2011)

A Methodological Proposal to Design a Trace-Based System to Qualify Cognitive Features Inscribed in Digital Learning Resources

Françoise Greffier and Federico Tajariol

"Objets et Usages Numériques" Research Team,
ELLIADD Laboratory,
University of Franche-Comté
{francoise.greffier,federico.tajariol}@univ-fcomte.fr

Abstract. To reuse existing digital learning resources, a teacher has to perform a complex and creative in order to adapt them to the students' needs and fit their cognitive abilities. Nevertheless, metadata do not generally describe the cognitive facets of a digital learning resource. We give here a partial account of a trace-based system in its current state of development to index digital learning resources according to the metadata describing their "cognitive features". By "cognitive feature" we mean the cognitive activities (e.g., reading, listening, body interactions, etc.) associated with different message formats (e.g., text, audio, interactive animations, etc.).

Keywords: trace-based system, cognitive styles, indexation of digital learning resources.

1 Introduction

Information and Communication Technologies for Education (ICTE) allow teachers to enrich the content of a Digital Learning Resource (DLR) with different formats: a text, an audio file, an image, an interactive animation, etc. So, for example, someone who is asked to read a text uses different abilities and skills than when (s)he observes a diagram illustrating the same content, and, because of his(her) cognitive style [1], (s)he better understands through the former cognitive activity than by the latter. In this way, the ICTE enable the personalization of the learning experience [2] [3], in order to fit the specific learners' cognitive abilities and skills. Nevertheless, a teacher runs into difficulties because of the scarcity of details about the cognitive features of a DLR. To tackle this problem, we seek to a) define metadata to describe the cognitive features associated with the presentation format of a DLR, b) realize a trace-based system (TBS) which enables to establish the metadata and to carry out a system to index the DLR.

2 The System to Index the DLR (SID)

In order to approach the cognitive side of a DLR, we use the Multiple Intelligence (MI) Theory [4], which has been successfully applied to the design of TEL

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 603–604, 2013.

environments [5]. According to the MI theory, the intelligence of a person is a set of abilities and skills (s)he develops to solve problems and process information. The MI theory suggests different sets of intelligence (e.g., kinaesthetic, linguistic, logical-mathematical, musical, visuo-spatial). These sets are related to some specific semiotic components. For example, a person who has a visuo-spatial intelligence is able to better think through semiotic components such as visual elements (e.g., images, graphs, cards, colours, etc.) and structures (e.g., patterns, diagrams, etc.) rather than speeches (linguistic intelligence). Our proposal, called SID, is a trace-based system (TBS), which collects and transforms elements [6]. The SID works as a parser: it analyses a DLR as input (e.g., a file such as a .ppt, .pdf, .doc) and it counts the semiotic components of the DLR, recorded in a database. We conceived two ontologies: the “Cognitive Style” (CSo) ontology, which describes the intelligences related to the MI theory and the “Semiotic Descriptors” (SDo) ontology, which models the semiotic components of a DLR. For instance, the item « image » is the main descriptor for the visuo-spatial intelligence. The parser, developed under Java 1.7.0_11, uses some API such as *iText* (for PDF format) and *Apache Poi*, which includes some methods used to identify the elements of a text and to recognize images. We developed different algorithms, e.g., to distinguish a coloured picture from a diagram by counting the rate of colours in the set of pixels. Further, the SID computes the tokens of each semiotic components and it gives a discrete value of the set of intelligence related to each specific semiotic component (e.g., percentage of sentences for the linguistic set), and it finally defines metadata to identify the cognitive styles of DLR. Then, a tool for automatic indexing will be added to the TBS. In fact, in spite of some ontology-based image system [7] and some metadata (e.g., in the standard LOM, we observed that the DLRs are seldom indexed in general.

References

1. Zhang, L.F., Sternberg, R.J., Rayner, S. (eds.): Handbook of Intellectual Styles: Preferences in Cognition, Learning, and Thinking. Springer, New York (2012)
2. Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.): The Adaptive Web: Methods and Strategies of Web Personalization. Springer, Heidelberg (2007)
3. Türker, M.A., Görgün, İ., Conlan, O.: The Challenge of Content Creation to facilitate Personalized eLearning Experiences. Int. J. E-Lear. 5(1), 11–17 (2006)
4. Gardner, H.: Frames of mind: The theory of multiple. In: Basic Books, New York (2004)
5. Tracey, M.W.: Design and development research: a model validation case. Ed. Tech. Res. Dev. 57, 553–571 (2009), doi:10.1007/s11423-007-9075-0
6. Settouti, L.S., Prié, Y., Marty, J.C., Mille, A.: A Trace-Based System for Technology-Enhanced Learning Systems Personalisation. ICALT, 93–97 (2009)
7. Hyvönen, E., Styrman, A., Saarela, S.: Ontology-Based Image Retrieval. In: Towards the Semantic Web and Web Services. In: Proceedings of XML Finland Conference, pp. 15–27 (2002)

Towards an Integrated Model of Teacher Inquiry into Student Learning, Learning Design and Learning Analytics

Cecilie Hansen[1], Valérie Emin[2], Barbara Wasson[3], Yishay Mor[4], María Jesús Rodríguez-Triana[5], Mihai Dascalu[6], Rebecca Ferguson[4], and Jean-Philippe Pernin[7]

[1] InterMedia, Uni Health, Uni Research, Bergen, Norway
cecilie.hansen@uni.no
[2] S2HEP, Institut français de l'Éducation, ENS-Lyon, Lyon, France
valerie.emin@ens-lyon.fr
[3] Department of Information Science and Media Studies, University of Bergen, Norway
barbara.wasson@uib.no
[4] Institute of Educational Technology, The Open University, UK
yishay.mor@open.ac.uk, rebecca.ferguson@open.ac.uk
[5] GSIC-EMIC, University of Valladolid, Spain
chus@gsic.uva.es
[6] Department of Computer Science, University Politehnica of Bucharest, Romania
mihai.dascalu@cs.pub.ro
[7] Laboratoire d'Informatique de Grenoble, Grenoble, France
jean-philippe.pernin@imag.fr

This poster introduces the first version of an integrated model of three traditions of research in TEL: Teacher Inquiry into Student Learning (TISL) [1], Learning Design (LD) [2] and Learning Analytics (LA) [3]. The integrated model, is based on four existing models: TISL Heart Model [4], Design Inquiry Model [2], Scenario Design Process Model [5], and the Model for Integrating Design and Analytics in Scripting for CSCL (MIDAS4CSCL) [6]. The result is leading towards a new strand of inquiry, called *teacher-led design inquiry of learning*.

TISL addresses the professional development of teacher practice by investigating student learning through action-oriented, evidence-based teacher-led research, with a particular focus on formative e-assessment. LD is the act of devising new practices, plans of activity, resources and tools aimed at achieving particular educational aims in a given situation, informed by subject knowledge, pedagogical theory, technological know-how, and practical experience. Although LA can be seen as "the measurement, collection, analysis and reporting of data about learners and their contexts" (LAK'11), it aims to extend beyond proposing tools responsible for analysing learning outcomes, providing a holistic, dynamic and formative view of learning processes.

Fig. 1 depicts the proposed model with emphasis on the target audiences its methods and tools specifically designed for practitioners: teachers who wants to inquire into the learning of their students, teachers/practitioners as designers of pedagogical scenarios, and teachers who want to monitor students' activities. We envisage this model to be used for designing better learning analytics tools, specifically tailored for learning scenarios. The model provides a context for these different fields to complement one another and build on each other's strengths.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 605–606, 2013.

TISL Heart	Design Inquiry Model	Scenario Design Model	MIDAS4CSCL	Integrated Model
Kick-off	Imagine	Idea of the learning scenario, intentions, and pedagogical approaches		Initiation
Set assumptions	Investigate		Context Analysis. Definition of prerequisites	Context analysis or investigation
Define R&D Question		Design of the scenario for the class/context, successive iterations	Define learning objectives	Formulation of the design objective and the research question
Design method to answer the question	Inspire and ideate		Select the pedagogical pattern. Configure the activity flow, groups, and resources.	Design method to achieve learning objectives and to answer research question(s)
Enact changed teaching and assessment	Prototype	Enactment and successive adjustments	Instantiate the design. Enact the design.	Enactment
Evaluate learning outcomes. Provide summative feedback.	Evaluate	Evaluation of the scenario enactment	Evaluate learning situation and design. Provide feedback.	Evaluation
Refine overall model (formative feedback loop)	Reflect	Reflection on the design, comments and patterns. Re-design and decontextualization	Re-design	Reflection and re-design

Fig. 1. The Integrated Model

The integrated model can be considered a promising direction for future development of educational practices, as well as a rich field for research. LD and LA are currently gaining ground as potent approaches to technology-enhanced educational practice. Yet, to gain validity, LD needs to incorporate data and to gain impact, whereas LA needs to influence design. Thus, both LD and LA can only manifest their full potential if they are integrated in a coherent cycle of inquiry, as through the TISL cycle and through innovation. We see the model proposed here as a first step in this direction.

References

1. Clark, W., Luckin, R., Jewitt, C.: Deliverable D5.1 Methods and Specifications for TISL Components V1. NEXT-TELL Consortium, EU (2011)
2. Mor, Y., Craft, B.: Learning design: reflections on a snapshot of the current land- scape. Research in Learning Technology 20 (2012)
3. Cooper, A.: A Brief History of Analytics. Analytics Series, vol. 1 (2012)
4. Hansen, C., Wasson, B.: The TISL Heart: A Model and Method for Teacher Inquiry into Student Learning (submitted)
5. Emin, V., Pernin, J.-P., Guéraud, V.: Model and tool to clarify intentions and strategies in learning scenarios design. In: Cress, U., Dimitrova, V., Specht, M. (eds.) EC-TEL 2009. LNCS, vol. 5794, pp. 462–476. Springer, Heidelberg (2009)
6. Rodríguez-Triana, M.J., Martínez-Monés, A., Asensio-Pérez, J.I., Dimitriadis, Y.: Towards a Monitoring-Aware Design Process for CSCL Scripts. In: Herskovic, V., Hoppe, H.U., Jansen, M., Ziegler, J. (eds.) CRIWG 2012. LNCS, vol. 7493, pp. 223–236. Springer, Heidelberg (2012)

Learning Instructional Design in a Project-Based, Technology-Enhanced Course

Ioana Hartescu

Lancaster University, UK
s.hartescu@lancs.ac.uk

Abstract. By using phenomenography, the research aims to explore the different conceptions of reality that various actors (undergraduate and postgraduate students, teachers, trainers and managers in placement organizations) have in the context of two project-based learning, technology-supported courses in "Instructional design of online courses" at the University of Bucharest.

Keywords: instructional design, project-based learning, technology enhanced learning, phenomenography.

1 Background

Building on the concepts of learning as being either deep learning (based on understanding) or surface learning (based on reproducing), Biggs [1] notes that when teachers use active teaching methods (such as project-based learning), students are encouraged to use a deeper approach to learning, while more passive methods (such as lectures) have no influence on their preferred approach.
Project-based learning (PBL) is a method that "organizes learning around projects" [2]. Being involved in authentic projects enables students to explore and apply the subject matter to complex situations, relevant to their future professional practice, developing skills such as accountability, independent learning, and collaboration [3]. Students are encouraged to plan their actions, make decisions, and work together in order to produce the desired outcome. In higher education, especially undergraduate, PBL is used in various disciplines such as international business [4] or marketing [5].

2 Research Design

For one graduate and one undergraduate course in Instructional design, I used PBL as a pedagogical approach, based on my experience of training instructional designers in professional contexts. Two cohorts of 22 students will design small e-learning courses for real organizations, placing the learning in an authentic work context. Students will meet with the clients' representatives to analyze the current situation and identify the learning needs; they will work with subject-matter experts; by using the free version of the CourseLab authoring tool the students will produce small online modules which they will present to the client to receive feedback. Their learning will be supported by an e-learning platform and they will also have regular weekly course meetings.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 607–608, 2013.

The proposed research aims to answer the following questions:

- What different experiences and perceptions students, teachers, and placement organizations staff have of PBL, the role of the teacher in PBL, and the role of technology in learning instructional design?
- How do students conceive of their own learning of instructional design?
- What are the implications of the findings for policy and practice?

Phenomenography, the chosen methodology, explores the various ways in which people experience the world; it aims to describe, analyze and understand these second-order perspectives called conceptions of reality [6]. Data from students will be collected using a two-level approach: the first level will include questionnaires, and the second level will include semi-structured interviews, focus groups, and online learning journals. This will ensure a wide enough spread of the various conceptions students have, while keeping the amount of interviews manageable. University staff and representatives of the placement organizations will also be interviewed.

3 Expected Results

Data collection ends in June 2013 and by September 2013 partial results will be available. The research will document the students' experience in using PBL in the context of a Romanian university, broadening the outcome space of the studied phenomenon.
The results can lead to further improvements in the way PBL is used in instructional design courses, and, by extension, in other related subjects, as other teachers become aware of the results. Although in many courses project reports account for a great part of the students' final grades, PBL is not frequently used, so, this method may extend to other disciplines, promoting interdisciplinary projects and an integrated curriculum. To my knowledge, this is the first time the Instructional Design course is offered in Romania at academic level. If successful, the course may become a permanent part of the curriculum and may even expand to a stand-alone degree. The professional training of instructional designers in the industry could also benefit from the results.

References

1. Biggs, J.: What the student does: teaching for enhanced learning. Higher Education Research & Development 31(1), 39–55 (2012)
2. Thomas, J.W.: A Review of Research on Project-Based Learning. Autodesk Foundation, San Rafael (2000)
3. Bell, S.: Project-based learning for the 21st century: skills for the future. The Clearing House 83(2), 39–43 (2010)
4. Korfhage Smith, R.: A Case Study in Project-Based Learning: An International Partnership. Journal of Teaching in International Business 21(3), 178–188 (2010)
5. Lopez, T.B., Lee, R.G.: Five principles for workable client-based projects: Lessons from the trenches. Journal of Marketing Education 27(2), 172–188 (2005)
6. Marton, F.: Phenomenography – describing conceptions of the world around us. Instructional Science 10, 177–200 (1981)

Enriching the Web for Vocabulary Learning

Róbert Horváth and Marián Šimko

Faculty of Informatics and Information Technologies,
Slovak University of Technology in Bratislava, Ilkovičova, 842 16, Bratislava, Slovakia
roberthorvath89@gmail.com, simko@fiit.stuba.sk

Abstract. People spend large amount of time browsing the Web while fulfilling various needs, but they find it difficult to spare some time for education. We believe that the time spent by browsing can be used more efficiently. We proposed a method for web augmentation during casual web browsing, which facilitates foreign language vocabulary learning. Our method substitutes words on a web page for their foreign equivalents and exposes a user to unknown words. We conducted a small qualitative experiment to study how users perceive word substitution in web pages and how it affects their foreign language vocabulary. Results showed that web page augmentation improves user's vocabulary size without negatively affecting his browsing experience.

Keywords: vocabulary acquisition, web augmentation, foreign language learning, computer-assisted language learning, NLP.

1 Web-Aided Vocabulary Learning

It has been shown that annotated reading has the positive effect on foreign language learning [2]. Users encounter foreign words in context, which is considered to be one of the key strategies for effective vocabulary acquisition [1]. Some approaches utilize close corpora and static reading systems [4], a few directly utilize web content [5], but we are not aware of any employing personalisation present in other web-based learning environments [3].

The Web represents a great opportunity for autonomous foreign language vocabulary acquisition. In our work, we propose a method for personalized web-aided foreign language vocabulary learning by augmenting web content (textual content). We employ a technique of word substitution, in which selected words in a native language are substituted with their foreign language translations. In general, the method does not restrict on the particular native-foreign language pair. The method selects terms appropriate for learning in the content of a web page a user chooses to read and replace them with their translations, so the user is exposed to mixed native and foreign language content. The method consists of the following main steps applied for every visited web page:

1. *Web Page Request.* A user makes a web page request in the browser.
2. *Web Page Augmentation.* We pre-process the web page to remove irrelevant information, translate the extracted content into the foreign language, select appropriate words/phrases for substitution and substitute them in the page.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 609–610, 2013.

3. *Web Page Reading and Interaction.* The user obtains the augmented web page. He accesses the page as intended, but is exposed to substituted words that he can click to reveal the original not-translated form.
4. *User Model Update.* Based on user's behavior on the augmented web page (time spent on a page, interaction with text, scrolling, etc.), the user model is being updated to keep the representation of user's knowledge accurate.

We created a web browser extension for Google Chrome Web browser to conduct a (controlled) qualitative experiment, in which we studied the influence of web augmentation on vocabulary learning and user experience. The study was conducted in a setting with Slovak and French as native and foreign language, respectively, involving ten participants with varying proficiency in French.

The major aim of our study was to find answers on research questions covering areas of user experience, learning performance, method parameters and browser extension usage. The participants were very excited about the approach. They fully understood the content of the augmented web pages. The substituted words were perceived as a minimal distraction, users enjoyed seeing them in general. Interestingly, everyone found effect on reading speed very reasonable and acceptable even if the measurements showed slowing down of 30 %. The experiment revealed that users are able to learn new vocabulary. They were able to remember and understand substituted words that occurred often in the text. The user responses indicate that multiple occurrences of the same word is the best way for most users to learn new vocabulary. The second place is held by the context of a word, which is helpful for both understanding and remembering the word. Some users mentioned context as the most important, but they considered it insufficient in general. The obtained results indicate that using our method we can improve user vocabulary without negatively affecting his experience.

The proposed approach has a great potential to attract foreign language learners as it decreases the cost of vocabulary acquisition. It better integrates TEL within our daily work by approaching the most popular medium: the Web.

Acknowledgments. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0675/11, the Slovak Research and Development Agency under the contract No. APVV-0208-10 and PeWe Group.

References

1. Jones, C.: Contextualise & personalise: Key strategies for vocabulary acquisition. RECALL-HULL 11(3), 34–40 (1999)
2. Laufer, B.: Electronic dictionaries and incidental vocabulary acquisition: Does technology make a difference? In: Proc. of EURALEX 2000, pp. 849–854 (2000)
3. Michlík, P., Bieliková, M.: Exercises recommending for limited time learning. Procedia Computer Science 1(2), 2821–2828 (2010)
4. Shei, C.-C.: FollowYou! an automatic language lesson generation system. Computer Assisted Language Learning 14(2), 129–144 (2001)
5. Streiter, O., Knapp, J., Voltmer, L., Zielinski, D.: Browsers for autonomous and contextualized language learning: tools and theories. In: IEEE 3rd Int. Conf. on Information Technology: Research and Education, ITRE 2005, pp. 343–347 (2005)

Digital Didactical Designs in iPad-Classrooms

Isa Jahnke, Lars Norqvist, and Andreas Olsson

Umeå University
Department of Applied Educational Science
Interactive Media and Learning (IML), Sweden
{isa.jahnke,lars.norqvist,andreas.olsson}@edusci.umu.se

Summary. Traditionally, Information and Communication Technology (ICT) "has been segregated from the normal teaching classroom" [12], e.g. in computer labs. This has been changed with the advent of smaller devices like iPads. There is a shift from separating ICT and education to co-located settings in which digital technology becomes part of the classroom. This paper presents the results from a study about exploring digital didactical designs using iPads applied by teachers in schools. Classroom observations and interviews in iPad-classrooms in Danish schools have been done with the aim to provide empirical evidence on the co-evolutionary design of both, didactical designs and iPads. The Danish community Odder has 7 schools where around 200 teachers and 2,000 students aged 6-16 use iPads in a 1:1 iPad-program. Three key aspects could be explored: The teachers' digital didactical designs embrace a) new learning goals where more than one correct answer exists, b) focus on producing knowledge in informal-in-formal learning spaces, c) making learning visible in different products (text, comics, podcasts etc.). The results show the necessity of rethinking traditional *Didaktik* towards Digital Didactics.

Digital Didactical Designs. Teaching is more than information delivery and remembering facts (surface levels). An appropriate teaching design enables a "conceptual change" to deepen learning [23] e.g., critical, self-reflections, multi-perspectives [19]. The goal of a *design* is to provide possibilities to enable learning [37]. The term *didactical design* is inspired by from the German concept of Didaktik by Klafki [24] [25] and follows Hudson [14] Fink [11] and Lund & Hauge [28] who stress the differences of teaching concepts and learning activities and call them designs for teaching and designs for learning. A didactical design includes the design of (a) teaching objectives, (b) learning activities (co-constructing knowledge), (c) different forms of process-based feedback, (d) social relations (dynamics of social roles, interactions), (e) ICT, mobile technology or tablets (as a booster for learning). An appropriate design is when these elements are visible in a "constructive alignment" [4]. According to the study by Bergström [3], process-based assessment is the most effective method to foster learning. The difference from an instructional design is that didactical designs also include the design of the social relations, e.g., student-teacher-interaction and peer interaction, "dynamics of social roles" [20] [14] (Figure 1).

The *digital* didactical design is the advanced model that integrates a fifth dimension: Educational Technology. To each of the design elements the design-question is, how can ICT and/or mobile devices like iPads support teaching aims and learning activities? We expect different degrees of the iPad-use: from a low, a medium to a

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 611–612, 2013.

high extent. The implementation of new technology in education means to rethink the existing underlying didactical concepts [18] [19]. ICT can play an important role in making learning visible. In two case studies, Mårell-Olsson & Hudson [29] illustrate different types of digital portfolios (traditional online applications for stationary computers or laptops) in which students develop the ability to "collect, organize, interpret and reflect on their own individual learning and practice, and become more active and creative in the development of knowledge" (p.73). An appropriate digital didactical design is defined as the "constructive alignment" of all five elements [4] increases the likelihood to enable surface and deeper learning. The research question is: What digital didactical designs do teachers apply, how and why?

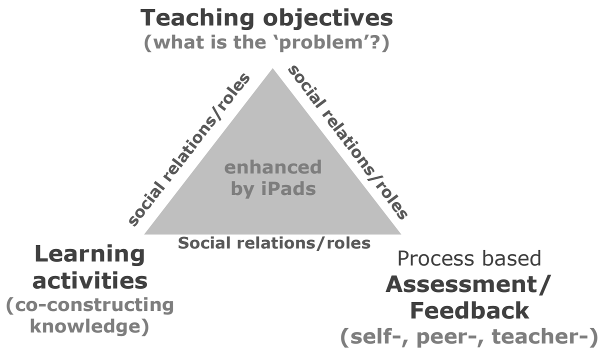


Fig. 1. Digital Didactical Design (DDD)

Findings. Part of a broader project, we studied 15 classrooms: A well-designed classroom correlates with a high extent of iPad use and vice versa, a not aligned didactical designed classroom correlates with a low extent of iPad use. Five patterns occurred:

A. Innovative iPad-classrooms: Alignment of didactics & technology (5 classes)
B. Almost innovative classrooms (not as strong as in pattern A) (3 classes)
C. Weak alignment of DDD but learning benefit through iPad-integration (1 class)
D. Potential for DDD, alignment differs, medium/low extent of iPad-use (4 classes)
E. Applied designs limited learning experiences, re-alignment of a digital didactical design is required; better without iPads? (2 classes)

We studied the innovative iPad-classrooms in detail and explored **key principles**:

- new type of learning goals where more than one correct answer exists,
- focusing on learning as process of co-producing, informal-*in*-formal spaces
- making school learning visible in different products (choice of tasks),
- innovative teachers used apps that are primarily not built for education.

These principles illustrate Digital Didactics do not exist of didactical designs *plus* ICT; new Digital Didactics Designs emerged to boost surface and deeper learning.

References. Due to limited space, the full list of reference is online available at www.isa-jahnke.com/publications (2013): Jahnke, I., Norqvist, L., & Olsson, A. (2013): Digital Didactical Designs in iPad-classrooms. In: Conference Proceedings of ECTEL 2013. Cyprus. Extended Online Material.

A Hybrid Multi-recommender System for a Teaching and Learning Community for the Dual System of Vocational Education and Training

Xi Kong[1], Susanne Boll[2], and Wilko Heuten[1]

[1] OFFIS – Institute for Information Technology, Oldenburg, Germany
{xi.kong,wilko.heuten}@offis.de
[2] Media Informatics and Multimedia System, University of Oldenburg, Germany
susanne.boll@informatik.uni-oldenburg.de

A number of recommender systems (RS) have been or are being developed in the context of technology enhanced learning (TEL). However, there seems to be a lack of research focusing on the dual system of vocational education and training (VET). The knowledge transfer and sharing in a dual system has its own particularities and difficulties because of its inborn nature. Firstly, in the dual system, apprentices in-company practice training is at the workplace and their theoretical education is in the classroom in a vocational school. The transfer of know-how takes place in two different geo-location. Secondly, different stakeholders are involved in the dual system in the sense of knowledge transfer and sharing. We consider three roles for analyzing the knowledge transfer and sharing, i.e. trainee, trainer and teacher. This suggests, six flows of knowledge transfer and sharing may occur, i.e. flow: teacher ⇔ trainee, trainer ⇔ trainee, teacher ⇔ trainer, trainee ⇔ trainee, trainer ⇔ trainer, teacher ⇔ teacher. Knowledge transfer and sharing among trainee, teacher and trainer is not easy, if there is no fitting instrument to support this process. The project *expertAzubi* addresses these issues to develop an online teaching and learning community using Web 2.0 technologies. We propose a hybrid multi-recommender approach to (1) support knowledge transfer and sharing among trainees, teachers and trainers in a pro-active way, and (2) use particularly the user generated contents to generate personalized recommendations to motivate learners to initialize an active lifelong learning.

The goals and tasks for which the RS are being used in general are for instance: find good items, recommend sequence [1]. In the TEL context the user tasks that could be supported by RS are e.g. find novel resources, and find peers [2]. Our RS covers multiple goals and tasks described in [1,2], i.e. find good items, just browsing, find novel resources, find peers. Furthermore we also emphasize other aspects, i.e. *Interest Degree* which aims at better supporting knowledge sharing, and combine the different types of recommenders. An initial focus group with apprentices shows that different forms or types of recommenders are wanted. We develop a hybrid multi-recommender system with three types of personalized recommenders, i.e. a hybrid multi-dimensional article recommender, a people recommender which finds potential peer learners or potential learners to build learning group, and a knowledge-based video recommender which offers another type of learning media to motivate learners.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 613–614, 2013.

In the *expertAzubi* community users may create, reference, cite or write learning materials. They may create tags, attach files, images or videos for the articles they created and rate or give comments to an article by other community members. Community members may publish an article to the whole community or certain groups. We analyzed the user generated contents from three initial pilot vocational school classes of the *expertAzubi* community. Based on the first analysis we developed and defined four dimensions, i.e. *Similarity Degree*, *Quality Degree*, *Interest Degree*, and *Diversity Degree*, for the recommendation process of the article recommender. *Similarity Degree* indicates how similar the content of a contribution in the community is to the current user profile. We aggregate the training reports of users as user profiles. *Quality Degree* measures the quality of a contribution based on the opinion of the community members. *Interest Degree* indicates to what extent a contribution draws the attention of the community members. *Diversity Degree* shows how broad or diverse the content of a contribution is. We then make a linear combination with weighting factors of these four dimensions to get a score for each contribution. This is the first recommendation list. Another two approaches are incorporated to achieve an overall RS. (1) We extract the most visited contribution of each user over all sessions from the log history. (2) Users may participate in different groups or start a new group in the community, e.g. school class, community members with similar learning goals. We extract the log history of the user's most often visited contributions in each group they belong to. Another two recommendation lists based on user behavior and user behavior in groups respectively are then generated. Finally, the hybrid article recommender follows the 'mixed' approaches of [3]. The recommendations from the three different components are combined based on their normalized scores before being presented to the users. Prediction accuracy, coverage and diversity were chosen as metrics to evaluate the hybrid article recommender due to our objectives of RS. The initial offline evaluation indicates that our algorithm can suggest users substantially useful items.

Acknowledgments. This research is funded by BMBF – Federal Ministry of Education and Research, Germany. Our project partners are CoC – Centers of Competence, ITB – Institute Technology and Education and TZI – Center for Computing and Communication Technologies. The authors would like to thank the students, Lennart Weller, Bengt Lüers, Renke Grunwald and Steffen Saß for their contributions to the development of the RS and the *expertAzubi* community.

References

1. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. 22, 5–53 (2004)
2. Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., Koper, R.: Recommender systems in technology enhanced learning. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook. Springer (2011)
3. Burke, R.: Hybrid web recommender systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web. LNCS, vol. 4321, pp. 377–408. Springer, Heidelberg (2007)

Reflective Learning in the Workplace: The Role of Emotion

Birgit R. Krogstie and Monica Divitini

Norwegian University of Science and Technology
{birgitkr,divitini}@idi.ntnu.no

Abstract. We present a model that refines current understanding of reflective learning in the workplace by relating emotion to the reflection process.

1 Background

In the workplace, work and reflection on work are intertwined and feed into each through what can be conceptualized as a reflective learning cycle [1].

Emotion is essential to work and reflection. For instance, positive affect generally leads to more creativity [2], and emotions can influence memory [3] and decision making [4]. People's repertoire of actions is affected by emotion, negative mood states biasing attention toward threatening features of the environment [5]. Handling emotions is crucial e.g. in service and healthcare, and through emotion work [6] people manage what they feel and what feelings they display. The basic process whereby people seek to redirect the flow of their emotions is called emotion regulation.

In technology enhanced learning recent years have seen a strong interest in emotions during learning [7], e.g. in affect-sensitive intelligent tutoring systems and open learner models. In reflective learning in the workplace, tools for capturing, presenting and sharing emotion-related information have been developed. Successful tool implementation however critically depends on understanding user requirements and constraints in the complex work setting, and in the case of reflective learning this includes understanding the role of emotion. Currently the literature does not offer a coherent framework for addressing emotion in reflective learning in the workplace. We developed a model to meet this need, drawing on a wide range of literature especially from psychology. The relevance of the model has been validated with empirical data from dementia care, collected through observation and in-depth interviews.

2 A Model Relating Emotion to the Reflective Learning Process

Our starting point is a four-staged reflective learning cycle [1]. We present an emotion view of the cycle outlining key relationships to emotion regulation for each stage.

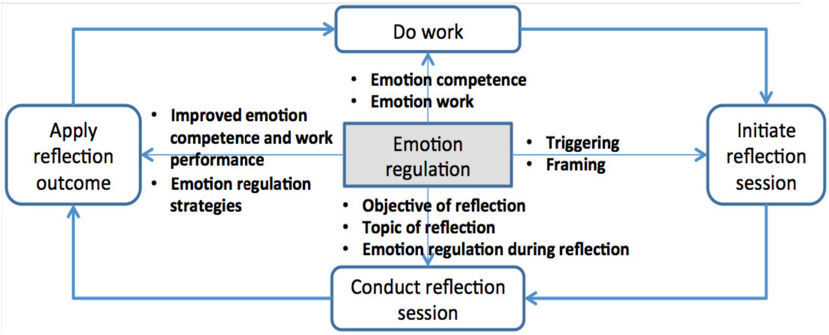
<u>During work</u>, handling emotionally demanding episodes through **emotion work** results in work experiences that one may need to reflect upon later. Individual abilities to handle emotions include **emotional competencies** that, crucially, can be improved through learning [8].

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 615–616, 2013.

For initiating reflection, emotions are an essential component of **triggering**, associated with a desire to get out of a negative emotional state, e.g. to resolve cognitive dissonance [5]. Emotion can also be an element in the **framing** of reflection, as a topic for reflection and/or as a condition for successfully conducting reflection.

When conducting reflection in a reflection session, reaching a desired emotional state (e.g. getting out of a negative state) can be an **objective of reflection**. Emotion can be a **topic of reflection**, e.g. when discussing emotion work. **Emotion regulation happening during reflection** is regarded as essential to reflective learning [9], e.g. as emotions originating in the experiences reflected upon have to be dealt with.

When applying a reflection outcome to work, **improved emotion competence and work performance** can be a type of resulting change to work. This includes **improved emotion regulation strategies**, e.g. in handling particular work situations.



In the design of computer-supported reflective learning solutions in the workplace this model can help systematically address emotional aspects of reflective learning.

References

1. Krogstie, B., Prilla, M., Pammer, V.: Understanding and Supporting Reflective Learning Processes in the Workplace: The CSRL Model. In: Proc. EC-TEL 2013, Vol. 8095, pp. 151–164 Springer, Heidelberg (2013)
2. Amabile, T.M., et al.: Affect and Creativity at Work. Adm. Sc. Q 50, 367–403 (2005)
3. Sharot, T., Phelps, E.A.: How arousal modulates memory: Disentangling the effects of attention and retention. Cognitive, Affective, & Behavioural Neuroscience 4(3) (2004)
4. Forgas, J.P.: Mood and Judgment: The Affect Infusion Model (AIM). Psychological Bulletin 117(1), 39–66 (1995)
5. Hudlicka, E.: To feel or not to feel: The role of affect in human-computer interaction. International Journal of Human-Computer Studies 59(1), 1–33 (2003)
6. Hochschild, A.R.: Emotion Work, Feeling Rules, and Social Structure. American Journal of Sociology 85(3), 551–576 (1979)
7. D'Mello, S.K., Calvo, R.A.: Significant Accomplishments, New Challenges, and New Perspectives. In: Calvo, R.A., D'Mello, S.K. (eds.) New Perspectives on Affect and Learning Technologies, pp. 255–271. Springer, New York (2011)
8. Nelis, D., et al.: Increasing emotional intelligence (How) is it possible? Personality and Individual Differences 47, 36–41 (2009)
9. Festinger, L.: A theory of cognitive dissonance. Stanford University Press (1957)
10. Boud, D., Keogh, R., Walker, D.: Reflection: Turning Experience into Learning. Routledge, Falmer (1985)

Design and Implementation of Conversational Agents for Harvesting Feedback in eLearning Systems

Karsten O. Lundqvist, Guy Pursey, and Shirley Williams

University of Reading, School of Systems Engineering,
Whiteknights, Reading, RG6 6AY, United Kingdom
k.o.lundqvist@reading.ac.uk

Abstract. Traditionally conversational interfaces, such as chatbots, have been created in two distinct ways. Either by using natural language parsing methods or by creating conversational trees that utilise the natural Zipf curve distribution of conversations using a tool like AIML. This work describes a hybrid method where conversational trees are developed for specific types of conversations, and then through the use of a bespoke scripting language, called OwlLang, domain knowledge is extracted from semantic web ontologies. New knowledge obtained through the conversations can also be stored in the ontologies allowing an evolving knowledge base. The paper describes two case studies where this method has been used to evaluate TEL by surveying users, firstly about the experience of using a learning management system and secondly about students' experiences of an intelligent tutor system within the I-TUTOR project.

Keywords: Conversational Interface, Chatbot, Ontology, AIML.

1 Summary

The primary aim of the research behind this paper is to develop a method for implementing conversational agents, also known as chatbots, which can change domain knowledge easily and expand the base of knowledge through conversation for specific conversational situations. This technology could potentially be used in many situations, however this system has exclusively been used within learning situations. By using the developed ontological language (Owllang) together with a traditional chatbot development language a survey chatbot was developed.

This system works by using language prediction. Human is not random but follows a Zipf curve distribution of possible responses to preceding utterances [1]. For instance, if person A says 'Can I try?' to person B, it is very unlikely that person B will respond '2 plus 2 is 4'. The chatbot developer, therefore, can try to cover as many user responses to each chatbot output as possible, using a descriptive chatbot language such as the Artificial Intelligence Meta Language (AIML) [2] to build a map of possible responses to human interaction, in this paper termed conversational trees. One of the earliest examples of this approach is ELIZA which utilized the basic structure of Rogerian therapy sessions to emulate a psychiatrist [3]. This is a trial and error approach where the developer constantly monitors the chat logs to see if new chat

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 617–618, 2013.

patterns have been discovered that need to be added to the conversational trees. The problem with this approach is that it is laborious to change any knowledge within the conversational trees. The developed system aids that by allowing knowledge injection into the conversational trees from standard ontologies that can be used to describe world knowledge [4].

A surveying robot has been created that questions users about their opinions. It works in English, and can respond to 8 different types of responses from the users. The system has been tested to evaluate TEL. The first case study was a surveying students about their user experience of the University's Blackboard system. The second test case study has been as an integrated surveyor within an Intelligent Tutoring System through the i-tutor project. Creating these new surveyor chat systems using an ontological approach has proven to be very easy compared to the natural language approach.

It could be argued that the chatbot just works as an automated interviewer, and a human evaluator still has to read through all of the logs, however the chatbot automatically stores what it infers from the communication using the knowledge in the ontologies. These ontologies can be examined using semantic web tools and statistical tools. What the human evaluator needs to do is extract the gist of the value-added information from within the chat, as this is information which would be extremely difficult to automate. The users' responses has also been analysed and the system has proved to understand most users feedback with a 96.7% success rate. There are only minor issues with double negatives and one user clearly misunderstanding the chatbot's question.

References

1. Zipf, G.K.: Selected Studies of the Principle of Relative Frequency in Language (1932)
2. Wallace, R.: AIML Overview (accessed September 10 2012)
3. Wiezenbaum, J.: ELIZA - A computer Program For the Study of Natural Language Communication Between Man and Machine. Communications of the ACM 9(1), 36–45 (1966)
4. Guarino, N., Giaretta, P.: Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: Towards Very Large Knowledge Bases, pp. 25–32. IOS Press (1995)

A Context Modelling System and Learning Tool for Context-Aware Recommender Systems

Christos Mettouris, Achilleas P. Achilleos, and George Angelos Papadopoulos

Department of Computer Science, University of Cyprus, Nicosia, Cyprus
1 University Avenue, P.O. Box 20537, CY-2109
{mettour,achilleas,george}@cs.ucy.ac.cy

Abstract. A critical contextual modelling issue in context-aware recommender systems research has to do with developing domain and application specific models that offer no reuse and sharing capabilities. Developers and researchers struggle to design their own models without any guidance, often resulting in overspecialized, inefficient and incomplete context models. On the basis of a prior work, we have developed an online context modelling system and learning tool that is able to guide and teach CARS developers and researchers through the process of CARS context modelling, on how to build their own context models in a way that offers sharing and reuse, while at the same time it teaches them important modern concepts derived by CARS research, significantly advancing as a result their knowledge in the field.

Keywords: Context Modelling Learning Tool, Context Modelling Framework, Context-Aware Recommender Systems, Context-Awareness.

Context-aware recommender systems (CARS) use context data to provide better and more personalized recommendations [1]. Context modelling is often used to model the contextual parameters to be used during the recommendation process. A critical contextual modelling issue in CARS is related to developing domain and application specific models that only represent information on the particular application domain (e.g. movies, restaurants nearby). Our review on recommender systems that use contextual and conceptual models [2] had revealed that most CARS and semantic recommenders in the literature are domain specific. Domain specific models cannot be applied in other domains, while application specific models cannot be applied to other recommenders even of the same application domain. By constructing domain and application specific contextual models, many different and very specific models are produced with no reuse and sharing capabilities. Moreover, developers and researchers struggle to design their own models as they think appropriate and according to their own knowledge and skills, with no reference model to use, no guidance and strictly based on the application at hand, often resulting in overspecialized, inefficient and incomplete contextual models.

We have attempted to address the above contextual modelling problem in a prior work by proposing a generic, abstracted contextual modelling framework for CARS, which developers and researchers can use for guidance through the process of properly defining the context for their application [2]. The modelling framework is

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 619–620, 2013.

essentially a model template designed and built as a UML class diagram by using the Eclipse EMF. The framework is thoroughly described in [2]. Although this framework was developed as a UML class diagram, it could be mainly used in a theoretical manner, as a schematic reference rather than a modeling tool: (i) it is not an easy and straightforward procedure for developers to extend or instantiate a UML class diagram in order to build their own context models, (ii) it is time consuming, (iii) it requires programming knowledge and skills and (iv) it does not offer guidance and learning of important concepts. Therefore, based on [2], we have developed an easy to use, easy to extend online web-based system called “Context Modelling System and Learning Tool” [3]. The tool, through appropriate interfaces and clickable components is able to guide and teach developers and researchers towards a more efficient, effective, easy and correct selection and usage of context properties for building their own application-oriented context models and context instances, allowing at the same time for sharing and reuse of context models and instances among applications, regardless of their domain. CARS developers are guided through the process of uniformly modelling the most important contextual parameters for their applications, as well as on how to use the modelling system to build and/or extend context models and instances in order to be applied for their own application. The system is also able to introduce to developers and new researchers modern concepts derived by CARS research that they might not be familiar with, such as the “context dependent rating data”, the “supposed context”, the “static/dynamic context”, the “context weights”, etc., as well as the role of such concepts in a context model and a recommendation process.

We are currently in the stage of evaluating our system by urging postgraduate students and researchers to use the system to learn about CARS and their concepts, build their own context models, applications and context instances and provide valuable feedback on both the learning process and the context model building process. Note that evaluators are not supported with lessons, guidelines or other information relevant with the system prior to their actual interaction with it. Based on this feedback, the context modelling system and learning tool is continuously updated and extended. As future work, an important goal is to enable CARS developers to attach recommendation algorithms to their context models in order to create CARS in an innovative and straightforward way that also supports sharing and reuse of context models.

References

1. Adomavicius, G., Tuzhilin, A.: Context-aware recommender systems. In: Proceedings of the 2008 ACM Conference on Recommender Systems, pp. 335–336 (2008)
2. Mettouris, C., Papadopoulos, G.A.: Contextual Modelling in Context-Aware Recommender Systems: a generic approach. In: Haller, A., Huang, G., Huang, Z., Paik, H.-Y., Sheng, Q.Z. (eds.) WISE 2011 and 2012 Combined Workshops. LNCS, vol. 7652, pp. 41–52. Springer, Heidelberg (2013)
3. CARS Context Modeling System and Learning Tool, `http://www.cs.ucy.ac.cy/~mettour/phd/CARSModeling/`

Sharing the Burden: Introducing Student-Centered Orchestration in Across-Spaces Learning Situations

Juan A. Muñoz-Cristóbal, Luis P. Prieto, Juan I. Asensio-Pérez, Iván M. Jorrín-Abellán, Alejandra Martínez-Monés, and Yannis Dimitriadis

GSIC-EMIC Group, Universidad de Valladolid, Spain
{juanmunoz,lprisan}@gsic.uva.es, {juaase,yannis}@tel.uva.es, ivanjo@pdg.uva.es, amartine@infor.uva.es

Abstract. We propose the notion of "learning buckets": Configurable containers of positioned learning artifacts, that teachers may include in learning designs, and that are filled during the enactment by students or teachers.

Keywords: Augmented reality, orchestration, ubiquitous learning, student-centered.

1 Sharing the Orchestration Load Using Learning Buckets

The orchestration [1] of learning situations involving multiple physical and virtual spaces is challenging for teachers. To aid the orchestration of such learning situations, related literature [2] proposes the use of authoring tools to translate pedagogical ideas to a computerized format, as well as augmented reality (AR) to access virtual artifacts from physical spaces. However, most of these proposals force students to follow a predefined learning design in which most details (e.g., tools to use or learning artifacts to produce) are specified *a priori*, thus reducing student autonomy during enactment. This "agency issue" is defined by some authors [3], as a clear challenge for practitioners in ubiquitous learning scenarios. Teacher-centered designs are asked to evolve to student-centered ones, where activities can be managed by the students themselves, thus helping the improvement of the learning experience with the decisions taken by them [4]. As an intermediate solution between a purely teacher-centered solution (e.g., everything is pre-defined in a learning design) and an entirely student-centered approach (i.e., students have total freedom), we propose *learning buckets* as an aid for teachers in orchestrating such learning situations. Learning buckets can be seen as learning artifacts containers, included by teachers in the learning design of across-spaces learning situations. Buckets are configured with constraints, defining what the students are allowed to do within them. During the enactment of a learning situation in different spaces, students can fill the bucket with learning artifacts (e.g., instances of Web 2.0 tools, Web documents, 3D models), which are tagged with properties of the space of interest (e.g., the geographical location or the association with a fiducial marker). Since some decisions about artifacts (e.g., type, number, place) and the operations over them (create, update, delete) are delegated to the students during enactment, part of the orchestration load is transferred from the teacher to the students.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 621–622, 2013.

Additionally, such decisions represent a more student-centered approach, favoring students' agency. We summarize both aspects under the term student-centered orchestration, because facilitating teachers the use of student-centered approaches may as well be a form of sharing the orchestration load with students.

The pedagogical opportunities learning buckets may offer can be illustrated through an across-spaces learning situation sample scenario aimed at fostering orienteering skills in K-12. The situation is composed of two inter-related activities that take place in the classroom and at a park, respectively. In the first one (to take place in the classroom using a VLE such as Moodle), students are divided into three groups. Each group has to prepare an orienteering route for another group, consisting in a sequence of quiz questions to answer, located in different places of the park. Students can use different Web 2.0 tools to create each question. In the second activity (to take place physically at the park), each group must find the questions of its corresponding route (created by another group) using a map, orienteering techniques and a tablet with an AR application to access to the questions.

The described situation poses two main challenges to the teacher: how to allow students to create their own learning artifacts using Web 2.0 tools as well as positioning them in a physical space (increased flexibility and more student-centered orchestration); and how to control that creation process using a set of restrictions defined by the teacher (maintaining the pedagogical intentions of the teacher expressed in the learning design). In this case, a learning bucket for each group may be defined at design time by the teacher, and configured with constraints (e.g., types and maximum number of artifacts that can be created as well as the permitted operations and positioning types). Such buckets could be embedded in the VLE, allowing students to use them during the enactment.

Acknowledgements. This research has been partially funded by the Spanish Projects TIN2011-28308-C03-02 and VA293A11-2, and the European Project 531262-LLP-2012-ES-KA3-KA3MP.

References

1. Prieto, L.P., Dlab, M.H., Gutiérrez, I., Abdulwahed, M., Balid, W.: Orchestrating technology enhanced learning: a literature review and a conceptual framework. International Journal of Technology Enhanced Learning 3(6), 583–598 (2011)
2. Klopfer, E., Sheldon, J., Perry, J., Rosenheck, L., Squire, K., Mathews, J., Shapiro, R.B., Coulter, B., Dunleavy, M.: Augmented reality games: place-based digital learning. In: Proceedings of the 9th International CSCL Conference, Hong Kong, China (July 2011)
3. Cope, B., Kalantzis, M.: Ubiquitous Learning: An Agenda for Educational Transformation. In: Ubiquitous Learning, pp. 3–14. University of Illinois Press (2010)
4. Hannafin, M.J., Land, S.M.: The foundations and assumptions of technology-enhanced student-centered learning environments. Instructional Science 25, 167–202 (1997)

Virtual Communities of Practice in Academia: Automated Analysis of Collaboration Based on the Social Knowledge-Building Model

Nicolae Nistor[1,2,*], Mihai Dascalu[3], Stefan Trausan-Matu[3], Dan Mihaila[3], Beate Baltes[2], and George Smeaton[2]

[1] Ludwig-Maximilians-Universität, Faculty of Psychology and Educational Sciences
Leopoldstr. 13, D-80802 München, Germany
nic.nistor@uni-muenchen.de
[2] Walden University, Richard W. Riley College of Education and Leadership
100 Washington Avenue South, Suite 900, Minneapolis, MN 55401, USA
{beate.baltes,george.smeaton}@waldenu.edu
[3] University Politehnica of Bucharest, Faculty of Automatic Control and Computers
Splaiul Independentei 313, RO-60042 Bucuresti, Romania
{mihai.dascalu,stefan.trausan,dan.mihaila}@cs.pub.ro

1 Rationale

Participation in communities of practice (CoPs) [1], either in face-to-face, or in technology-based environments, leads to the accumulation of experience, stimulates the social construction of knowledge and the development of expertise. An overview over the occurring complex long-term processes may provide means to foster collaboration and social knowledge construction in CoP. In technology-based environments, however, this requires valid and reliable automatic analysis procedures of the community dialogue. In previous work [2], social knowledge-building was analyzed as a cohesion-based collaboration assessment model applied to individual chats. This study extends the procedure after aggregating discussion threads from asynchronous discussion forums and performs a validation by comparison with the results of a manual content analysis based on a critical thinking framework [3]. The developed tool makes headway on applying learning analytics in field research on virtual CoPs.

2 Tool Development

For the automated analysis of dialogue in asynchronous, text-based discussion forums, the messages were counted along with their authors' identifiers, then social network analysis was applied to extract betweenness centrality. As a representation of the underlying discourse, a cohesion graph was built for each forum discussion thread [2]. This was a multilayered graph with three types of nodes: (1) a central node, the overall discussion thread, (2) participant interventions and (3) sentences, the main

* Corresponding author.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 623–624, 2013.

units of analysis as forum posts usually integrate multiple phrases, on which the initial natural language processing pipe was applied [2]. As specificity of the previous graph, the cohesion expressing the strength of the links was viewed in terms of: (1) the *distance* between interventions; (2) *lexical proximity* reflected in identical lemmas and semantic distances within WordNet; (3) *semantic similarity* from Latent Semantic Analysis (LSA) vector spaces and Latent Dirichlet Allocation (LDA) topic models [2]. Links within the graph were established through the explicit referencing facility made available in the forum environment (the "reply to" links) or as automatically identified cohesive links. Afterwards, each intervention was automatically scored and the values are later on cumulated per vCoP member. The *importance score* integrates both a quantitative and a qualitative dimension. From a quantitative perspective, different Information Retrieval techniques were used, whereas quality was reflected in the topics coverage and in the cohesive links that were used to augment the local importance of each intervention. Eventually, all discussion threads were automatically aggregated, and a global social network [2] was built with all the involved vCoP members. On the other hand, *collaboration* as a complement to *participation* is assessed as a social knowledge-building effect reflected in the information transfer between different participants, by using the previous utterance scores and the corresponding cohesive links [2].

3 Tool Validation

The validation study was conducted in an academic vCoP emerging at a US American online university. The entire vCoP consisted of N = approx. 500 faculty members. During 23 months, they made a total of 7370 interventions. Applying the automated content analysis procedure, the overall social knowledge-building scores varied between 0 and 5927 ($M = 229.83$, $SD = 576.36$). To validate the automated analysis, a sample of 414 interventions made by $n = 15$ vCoP participants was manually analyzed applying the critical thinking framework developed and validated for assessing critical thinking in online discussions by Weltzer-Ward, Baltes and Knight-Lynn [3]. A strong bivariate correlation of $r = .72$ ($p < .001$) was found between the manually determined messages' relevance and the automatically determined importance scores, which suggests that the automated dialogue analysis provides results similar to the manual analysis based on the critical thinking framework.

Acknowledgment. This research was partially supported by the 264207 ERRIC–Empowering Romanian Research on Intelligent Information Technologies/FP7-REGPOT-2010-1 project.

References

1. Lave, J., Wenger, E.: Situated learning: Legitimate peripheral participation. Cambridge University Press, Cambridge (1991)
2. Dascalu, M., Trausan-Matu, S., Dessus, P.: Cohesion-based analysis of CSCL conversations: Holistic and individual perspectives. In: CSCL 2013, Madison, USA (2013)
3. Weltzer-Ward, L., Baltes, B., Knight Lynn, L.: Assessing quality of critical thought in online discussion. Campus-Wide Information Systems 26(3), 168–177 (2009)

Transforming the Campus into a Digitally Augmented Learning Space Using *etiquetAR*

Mar Pérez-Sanagustín[1], Javier Melero[2], Davinia Hernández-Leo[2], Carlos Delgado Kloos[1], and Josep Blat[2]

[1] Dpto. de Ingeniería Telemática, Universidad Carlos III de Madrid, Avd. Universidad 30, 28911 Leganés, Madrid
[2] ICT Department, Universitat Pompeu Fabra, C/Roc Boronat 138, 08018 Barcelona, Spain
`{mmpsanag,cdk}@it.uc3m.es`, `{javier.melero,davinia.hernandez}@upf.edu`

Abstract. *etiquetAR* is a web-authoring application designed towards learning purposes to support practitioners in the design of activities based on QR codes. This poster illustrates how *etiquetAR* QR codes were used to augment a University Campus as a support for a gamified tag-based learning activity. The QR codes generated change its behavior depending on the profile of the student interacting with the tag. The results show the successful application of *etiquetAR* in a real educational context.

Keywords: Augmented spaces, mobile learning, QR tags, personalization, game, user study.

1 Augmenting the University Campus with etiquetAR

Tagging technologies such as QR codes offer outstanding possibilities for transforming spaces into digitally augmented learning spaces. Tags, when attached to particular locations, add a digital layer of information that transforms the physical surrounding extending users' learning experience [1]. Although there are tools for generating QR codes, to the best of our knowledge, there are no tools oriented towards learning purposes. etiquetAR[1] [2] has been designed towards educational purposes with the following functionalities: (1) creating, editing and deleting QR codes with one resource per code, (2) associating profiles to each of these resources and (3) organizing codes into collections. The codes can be accessed using any QR reader installed in the students' smartphone and commented if necessary. Teachers can use their smartphones to delete and even hide an eventual inappropriate comment during the activity, on runtime.

etiquetAR was used to generate 17 QR codes for augmenting the Campus of the University Pompeu Fabra (UPF). Two teachers designed these codes as a support for a learning activity in which first-year undergraduate students registered to the Engineering Bachelor's degrees in Informatics, Telematics and Audiovisual Systems were

---

[1] etiquetAR Website: `http://etiquetar.com.es`

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 625–626, 2013.

invited to learn about the campus services and areas. The codes contained three questions, each associated to one profile corresponding to one of the three Bachelor's Degrees. Since each Bachelor Degree follows a curriculum focused on different areas of ICT, it was especially interesting to show the students distinct information when accessing to the tags. Also, the tags were designed so as to let students to add comments and see comments left by others on-the-fly, since the tags are dynamically updated on runtime with students' contributions. Once designed, the teachers distributed and attached the tags to 17 locations along the Campus areas.

2 A Gamified Tag-Based Learning Activity

The activity was prepared for the first week of the course "Introduction to Information and Communication Technology" (IICT). In this activity, students had to explore the campus looking for the QR codes and solving the different questions contained. A mobile application was implemented specifically for the activity to offer a gamified environment for the students. The mobile application was used to guide the students along the activity flow that required the interaction with the different QR codes. This application also included a QR reader and was in charge of producing log files registering the students' interactions with the tags. The computational representation of the game followed a conceptual model or defining educational mini-games [1].

The results show that etiquetAR was successfully used by the teachers and very well adopted by the students in this educational context. First, the teachers agree that etiquetAR is a usable tool and a good support for designing and enacting tag-based interactive activities. One of the teachers says: "*eitquetAR is a useful and simple tool that allows you to create in few steps QR-codes, edit and associate them to different resources and profiles*". Second, teachers value the profile functionality and the log files collected indicate that most of students accessed only the information related with their course profile. Third, students agree that the activity helped them to locate themselves in the campus while learning about the different areas and services. Also, teachers indicate that QR codes are especially interesting in "*closed controlled spaces where the technology GPS does no work*" [Teacher A] and "*in situations in which you expect the user to voluntarily access the information in the tag*" [Teacher B].

References

1. Rouillard, J., Laroussi, M.: PerZoovasive: contextual pervasive QR codes as tool to provide an adaptive learning support. In: 5th Interantional Conference on Soft Computing as transdisciplinar science and Technology, pp. 542–548 (2008)
2. Pérez-Sanagustín, M., Martínez, A., Delgado Kloos, C.: etiquetAR: tagging learning experiences, In: Hernández-Leo, D., Ley, T., Klamma, R., Papadopoulos, G. (eds.) EC-TEL 2013. LNCS, Vol. 8095, pp.573–576, Springer, Heidelberg (2013)
3. Melero, J., Hernández-Leo, D., Blat, J.: Towards the Support of Scaffolding in Customizable Puzzle-based Learning Games. In: Proceedings of the 11th International Conference on Computational Science and its Applications (ICCSA), Santander, Spain, pp. 254–257 (2011)

Detecting Implicit References in Chats Using Semantics

Traian Rebedea and Gabriel-Marius Gutu

University Politehnica of Bucharest, Department of Computer Science and Engineering,
313 Splaiul Independetei, Bucharest, Romania
traian.rebedea@cs.pub.ro, gabi.gutu@gmail.com

Abstract. Chat conversations with multiple participants are widely used in solving a wide range of CSCL tasks. One of the reasons for their success is that they encourage multiple conversation threads to exist in parallel, thus allowing multiple topics and ideas to be debated at the same time. These threads may be detected more easily if we would be able to identify the links that exist between the utterances of a conversation. This paper tries to explain whether semantic similarity measures from Natural Language Processing (NLP) may be successfully used to detect the links between utterances in CSCL chat conversations.

Keywords: Chats, CSCL, NLP, Semantics, Implicit Links.

1 Introduction

Chat conversations are used in a large range of Computer-Supported Collaborative Learning (CSCL) activities, especially for debating and solving difficult problems [1]. One of the reasons for the successful use of chats in CSCL tasks is because they allow the existence of parallel discussion threads that are inter-animating [2].

Discourse analysis does not provide a theory for processing multi-party conversation chats. However, there are new theories that propose the use of conversation (or coherence) graphs for chat analysis [2]. At the base of these theories is the existence of a multitude of links between utterances that might explain the evolution of the discussion threads. However, the discovery of these implicit links is a difficult task, mainly because of the different types of links that may arise in a conversation [3].

2 Are Semantic Links Useful for Analyzing CSCL Chats?

Semantic relatedness between proximal words is an important feature in most coherent discourses and various linguistic methods have been proposed. Some of them are called "*strong*" semantics (or *knowledge-based*) methods because they use lexical resources built by linguists, while others are called "*weak*" semantics methods and they only require a large volume of discourse corpora to compute the relatedness between any two words using statistics. Thus they are known as *corpus-based* methods.

Starting from a corpus of multi-party chats with explicit links added by participants and containing over 2500 utterances, a first experiment was to compare the distribution of the explicit links to the semantic similarity between the current utterance and

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 627–628, 2013.

utterances that are at a certain distance before it. Fig. 1.a shows the average semantic relatedness scores computed using Latent Semantic Analysis (LSA) [4]. There is a very good resemblance with the distribution of the explicit links in Fig. 1.b, with two important differences: (1) the semantic relatedness scores are varying with the distance less abruptly, and (2) the semantic similarity is decreasing constantly with the distance until it stabilizes at *dist=8*; afterwards the variations are so slightly that they are probably caused by the overall discourse coherence.

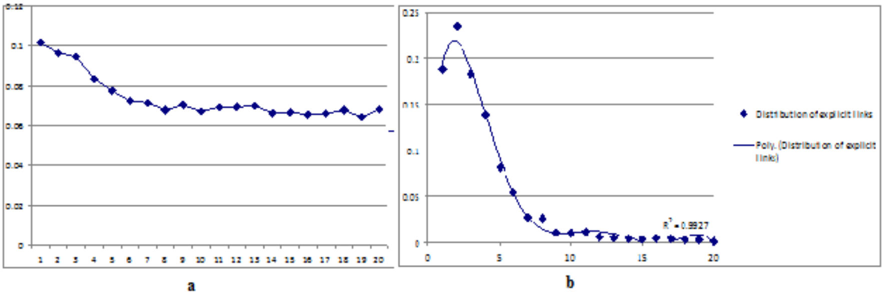


Fig. 1. (a) Average semantic relatedness computed using LSA, (b) Distribution of explicit links

On the other hand, it is clear that not all links can be identified using just the semantic similarity scores, especially the ones that depend on pragmatic and conversation-specific elements (e.g. adjacency pairs, signaling turns and others).

3 Conclusions

The distribution of explicit links is very similar to that of the semantic similarity scores, thus proving that both links and semantic similarity measures have a similar behavior on a global scale. However this is not always true when analyzing local elements: an individual link between two utterances is not always explained by semantic similarity and our studies show that only less than half of the links may be explained by semantic similarity or relatedness scores (knowledge and corpus-based) alone.

References

1. Stahl, G.: Studying Virtual Math Teams. Springer, New York (2009)
2. Trausan-Matu, S., Rebedea, T.: Polyphonic inter-animation of voices in VMT. In: Stahl, G. (ed.) Studying Virtual Math Teams, pp. 451–473. Springer, New York (2009)
3. Rebedea, T., Trausan-Matu, S., Chiru, C.: Inter-animation between Utterances in Collaborative Chat Conversations. In: Lim, H.L., Sudweeks, F. (eds.) Innovative Methods and Technologies for Electronic Discourse Analysis. IGI Global (2013)
4. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by latent semantic analysis. Journal of the American Society for Information Science 41, 391–407 (1990)

Learning with E-Flashcards – Does It Matter?

Inga Saatz and Andrea Kienle

University of Applied Science Dortmund
Germany
{Inga.Saatz,Andrea.Kienle}@fh-dortmund.de

Abstract. E-Flashcards are an approach to support students for active learning and individual feedback – even in large scale university courses. This paper presents results of a pilot study for using e-flashcards in two courses. The study shows a wide usage and acceptance of the e-flashcards in combination with further incentives to enhance user activity.

Keywords: active learning, e-flashcards, evaluation.

1 Introduction

The approach of using electronic (e)-flashcards in large-scale courses supports active and self-directed learning of the students and helps to deepen the knowledge on terms and facts (see [1, 2]). This paper presents a study, which investigates the user acceptance and the use of e-flashcards within large scale university courses. The study was conducted in the winter term 2012/13 in the two courses "Fundamentals of Business Computer Science" (1st semester) and "Database Management Systems 1" (3rd semester) at the Faculty of Computer Science, University of Applied Science and Arts, Dortmund (Germany). As an additional incentive for learner activation, the creation and rating of e-flashcards was awarded by bonus points which counted on the final examination mark.

The aim of the study was to get insights about the acceptance and usage of e-flashcards in large scale courses. Methods of evaluation were a logfile analysis of the e-flashcard server and a survey which gives answers to the question of the acceptance and difficulties during the usage of the e-flashcards.

2 Experiences and Results

In the first course "Fundamentals of Business Computer Science 1", 87 students registered on the of e-flashcard platform. During the study 61 participants created (= 70% of total group) 642 cards. 75 participants (86%) generated total of 1809 reviews and the instructor created 303 reviews additionally. The average of the created e-flashcards is 10 cards per author, which is closely linked to the organizational framework (0.5 points per created e-flashcard, max. 5 bonus points overall).

In the second course "Database Management Systems 1" 34 students registered on the e-flashcard platform. During the survey period, 26 participants (= 76% of total

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 629–630, 2013.

group) created 369 cards. 20 participants (59%) rated 602 e-flashcards. As in the other course, the number of users increased strongly in the last two weeks in which bonus points for the creation of e-flashcards could be achieved. In this course, 20 users (83%) created exactly 16 e-flashcards. This is linked to the organizational conditions directly (0.25 points per card, max. 4 bonus points overall).

The survey of the user acceptance of e-flashcard system was based on a questionnaire and was conducted in the first course. In this survey attended about 50% of the total participants (37 people). The technical learning support by means of e-flashcards was evaluated with good or very good by more than 80% of the participants. About 70% of the respondents indicated the creation of e-flashcards as a good method to deepen the course content. 32 of the 37 respondents (= 86%) quoted to have dealt with a learning topic before they created the corresponding e-flashcards. In addition, about 63% of the respondents said that they could remember the content better due to the creation of e-flashcards. By 74% of the respondents was indicated that the rating and commenting of e-flashcards is a good method for individual feedback on the created content. 60% of the respondents considered the ratings done by the lecturer as more important than those of other students.

Overall, the feedback regarding the user acceptance of e-flashcard system is positive. However, this high level of user acceptance is not reflected in the actual use of e-learning cards during the examination preparation phase. In this phase no additional incentives were given to increase learner activation. This resulted in a significantly lower use of the learning system compared with the use together with additional incentive in both courses.

3 Conclusion and Outlook

In this paper evaluation results of the use e-flashcards within two university courses were presented. This study showed that learners expressed a high level of acceptance of e-flashcards. However, a discrepancy occurred between the high user acceptance and the low use of the e-flashcard system during the examination preparation phase. This discrepancy is topic of our further research.

Acknowledgement. Thanks to Markus Förster and Robert Zimmer for support in evaluation.

References

1. Browder, B.M., Maura, L.R.: Guidelines for flash card instruction. Journal of Behavioral Education 3, 235–245 (1993)
2. Saatz, I., Kienle, A.: Increasing Quality in large scale University Courses. eleed, vol. 9 (last access: June 7, 2013), `http://eleed.campussource.de/archive/9/3655`

Interdisciplinary Knowledge Creation in Technology – Enhanced Learning

Eileen Scanlon[1], Grainne Conole[2], Gill Clough[1], and Canan Blake[1]

[1] The Open University, Walton Hall, Milton Keynes, UK
{eileen.scanlon,gill.clough,canan.blake}@open.ac.uk
[2] University of Leicester, Leicester, UK
grainne.conole@le.ac.uk

Abstract. The impact of the Internet on working practices has been profound, in terms of how people communicate, collaborate and network. In parallel, there has being increasing prominence given to interdisciplinarity as a means of addressing cross-disciplinary research challenges. This poster explores how interdisciplinary research can make better use of new technologies as a means of developing shared understanding. Interdisciplinary projects investigating Technology-Enhanced Learning (TEL) make a particularly relevant site for such research. We have found that a key means of support for the development of work on interdisciplinary projects is the development of mediating artefacts to support the articulation and process of discourse.

Keywords: Interdisciplinarity, Technology-Enhanced Learning, mediating artefacts.

1 Introduction

Technology offers an obvious way to break down the disciplinary boundaries in traditional academic practice because it is the medium through which research findings are translated into cultural products. As Moran [1] notes referenced in Conole et al. [4] '*both material objects and narrative devices, technologies have as much to do with re-imagining and presenting normative accounts of society as they do with providing local solutions to practical problems.'*

In interdisciplinary working a number of strategies can be adopted. One key aspect is the need to ensure that there is effective communication. The different perspectives amongst team members need to be articulated and interrogated in light of the research question being addressed. An ongoing iterative process of dialogic engagement and critical reflection is needed, so that the team can come to some degree of shared understanding and consensus. The time and effort needed to achieve this should not be underestimated. Technologies have the potential to act as powerful mediating-artefacts (Conole et al. [2]) in this process, by providing mechanisms for sharing and documenting understanding; and hence co-construction of new knowledge.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 631–632, 2013.

2 Mediating Artefacts

We illustrate this process in terms of a number mediating artefacts: a shared vocabulary wiki from the PI project (http://www.pi-project.ac.uk), and design and evaluation framework from the xDelia project (http:// www.xdelia.org). The projects share some similarities: they both involve interdisciplinary teams, and they both involve the design of a product, in the case of the PI project a toolkit to support inquiry and in the xDelia project, games to support the development of self-regulation to improve financial decision making.

In the PI project we were particularly interested in the use of mediating artefacts during the design phase. The development of the vocabulary on a shared wiki was very influential to communication in the project design phase. The first version of this project glossary wiki was produced during a workshop event. Some of the terms, which triggered this development included ‘scenario’, ‘activity’, ‘stages’, ‘phases’, ‘workflow’ and ‘script’. The xDelia project was European-funded, involving seven partner institutions working together in the attempt to explore the impact of emotion regulation on financial decision-making and address this by the use of serious games. We were exploring the activities of traders and investors led by finance experts and the effects of emotion biases on members of the public led by experts in personal financial capability. Also the team contained serious games researchers, researchers in sensor technology, and cognitive neuro-scientists and educational technologists. The key mediating artefact was the Design and Evaluation (D&E) framework [3]. This pictorial representation of the activity of design and evaluation and the relation between these activities was derived from a literature review. The D&E framework both acted as a model for the design of effective project interventions, providing structure and support for good practice and acted as a lens through which to reflect on what happened during the intervention, involving the stakeholders as reflective evaluators and feeding the findings back into the project on an ongoing basis.

In an interview study of some TEL researchers, clear and ongoing communication emerged as a key strategy for interdisciplinary working [4].

References

1. Moran, J.: Interdisciplinarity, 2nd edn. Routledge, London (2010)
2. Conole, G.: Capturing practice: the role of mediating artefacts in learning design. In: Lockyer, L., Bennett, S., Agostinho, S., Harper, B. (eds.) Handbook of Research on Learning Design and Learning Objects: Issues, Applications and Technologies, pp. 187–207. IGI Global, Hersey (2008)
3. Clough, G., Conole, G., Scanlon, E.: Behavioural finance and immersive games: a pan-European framework for design and evaluation. In: Same Places, Different Spaces - Proceedings ASCILITE, December 6-9, Auckland, New Zealand (2009)
4. Conole, G., Scanlon, E., Mundin, P., Farrow, R.: A commentary on interdisciplinarity. Prepared for the TLRP-TEL Programme (2010)

Evaluating System Functionality in Social Personalized Adaptive E-Learning Systems

Lei Shi, Malik Shahzad K. Awan, and Alexandra I. Cristea

Department of Computer Science, University of Warwick
{lei.shi,malik,acristea}@dcs.warwick.ac.uk

Abstract. Along with the theoretical and practical research on introducing a social dimension to adaptive educational hypermedia, the evaluation of such systems becomes more important. Existing evaluation methods are mostly based on statistical and qualitative analysis, in which researcher bias is built in and unavoidable. Moreover, they adopt either a traditional "as a whole" approach making it difficult to evaluate a system from different perspectives, or a "goal specified" approach, which only covers a specific aspect. Therefore, this study proposes a *generic method for evaluating system functionality*.

Keywords: adaptive educational hypermedia system, evaluation, framework.

1 Introduction

The evaluation of social personalized adaptive e-learning systems has been considered a complicated and effort-consuming task. Several evaluation approaches [1-3] have been developed, only covering limited perspectives.

2 System Functionality Evaluation Methodology

We have developed a *component-based evaluation framework*, which uses a Likert Scale (1-5) and has four components: *(1) System Functionality; (2) Learning Perspective; (3) System Prospect* and (4) *Overall System Classification*. The first one, System Functionality, is further detailed here. This component aims at using a Likert Scale to evaluate system functionalities from different performance aspects such as accessibility, effectiveness, operability, reliability, scalability and usability. We associate a weight, 'w' with each considered performance aspect, which represents its significance. The score value of this component, '$Comp_{FUNC}$', is calculated using Eq. 1 for representing the overall system value against the considered system functionalities. The score value is measured by taking summation of product of the considered performance aspects of system functionalities, '$SubSys_{(aspectID,subSysID)}$', and the associated weight, '$w_{(aspectID,subSysID)}$', where '*aspectID*' represents a considered performance aspect; '*subSysID*' represents a considered sub-systems; 'm' represents the number of the considered sub-systems. The generalized description of '$SubSys_{(aspectID,subSysID)}$' for each considered system functionality represented as

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 633–634, 2013.

'$F_{(aspectID,funcID,subSysID)}$', could be calculated using Eq. 2, where '*funcID*' represents a considered system functionality; '$w_{(aspectID,funcID)}$' represents the corresponding associated weight; 'n' represents the number of the considered system functionalities within the sub-system. '$F_{(aspectID,funcID,subSysID)}$' could be calculated using Eq. 3, where '$q_{(i,\ j)}$' represents the j^{th} question related to the considered system functionality, '$F_{(aspectID,funcID,subSysID)}$', answered by the i^{th} respondent; 'w_j' represents the corresponding associated weight of this question; 'k' represents the number of the questions related to the considered performance aspect of a considered system functionality; 'a' represents the total number of respondents. The term '1/a' is used to minimize the level of biasedness arising from the answers of a respondent.

$$Comp_{FUNC} = \sum_{subSysID\,=\,1}^{m} SubSys_{(aspectID,subSysID)} \times w_{(aspectID,subSysID)} \quad (1)$$

$$SubSys_{(aspectID,subSysID)} = \sum_{funcID\,=\,1}^{n} F_{(aspectID,funcID,subSysID)} \times w_{(aspectID,funcID)} \quad (2)$$

$$F_{(aspectID,funcID,subSysID)} = \frac{1}{a} \times \sum_{i=1}^{a} \sum_{j=1}^{k} q_{(i,j)} \times w_j \quad (3)$$

3 Conclusion

We have developed *component-based framework* for evaluating social personalized adaptive e-learning systems, aiming to reduce researcher bias, as well as provide a broader performance perspective of such a system. This evaluation methodology were used to evaluate the Topolor system [4] and allowed identifying a set of important features for further improvement [5]. Here, we have only described the first aspect of this framework, the System Functionality evaluation methodology.

References

1. Ozkan, S., Koseler, R.: Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. Computers & Education 53(4), 1285–1296 (2009)
2. Chao, R.J., Chen, Y.H.: Evaluation of the criteria and effectiveness of distance e-learning with consistent fuzzy preference relations. Expert Systems with Applications 36(7), 10657–10662 (2009)
3. Paechter, M., Maier, B., Macher, D.: Students' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction. Computers & Education 54(1), 222–229 (2010)
4. Shi, L., Al Qudah, D., Qaffas, A., Cristea, A.I.: Topolor: A Social Personalized Adaptive E-Learning System. In: Carberry, S., Weibelzahl, S., Micarelli, A., Semeraro, G. (eds.) UMAP 2013. LNCS, vol. 7899, pp. 338–340. Springer, Heidelberg (2013)
5. Shi, L., et al.: Evaluation of Social Interaction Features in Topolor - A Social Personalized Adaptive E-Learning System. In: The 13th IEEE International Conference on Advanced Learning Technologies (2013)

JACK Revisited: Scaling Up in Multiple Dimensions

Michael Striewe and Michael Goedicke

paluno - The Ruhr Institute for Software Technology
University of Duisburg-Essen, Germany

Abstract. In 2009 the authors of this paper published their proposal for a modular software architecture for computer aided assessments and automated marking. Since then, four more years of experience have passed. This paper reports on technical and organizational aspects of using the proposed architecture and the actual system in various scenarios.

1 Introduction

In recent years, several systems for automated grading and computer aided assessment haven been published [1]. As others, the authors of this paper published their architecture for a system called JACK in 2009 [2]. The system has been in continious use for four years since then and had to scale up to fulfil its duty in multiple dimensions: It had to support more users, a broader range of exercises, and more feedback types. At the same time, it turned from an experimental project to a sustainable part of the didactical concepts of the introductory course to programming at our university. There is a growing portfolio of exercises, which reduces setup time for each future term. It also makes it easier for new teachers to start working with JACK, so that continious use can be assured. Moreover, the quality of feedback messages for particular exercises could be improved over the years, making the system more helpful to the students.

2 Dimensions of Scaling

JACK has been in use for introductory courses to programming with up to 2'000 submissions per single programming exercise, but also for smaller project courses with more complx exercises, and for courses on mathematics. This implies scaling in various aspects: Number of submissions, size and complexity of exercises, and amount of feedback.

Scaling the number of submissions is the easiest problem that could be solved by using parallel grading components. The architecture of JACK considered this solution in its design, so applying it was mainly a matter of available server capacity. Each grading component is currently run on a virtual server, where a single CPU core with 2.5 GHz and 4 GB RAM is sufficient.

Scaling the size and complexity of exercises turned out to be no problem at all for automated grading. JACK was generally able to handle complex exercises

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 635–636, 2013.

involving libraries and multiple submitted files as quick as simple exercises involving just a single source code file. However, the setup time for a teacher to configure a complex exercise is of course larger than for simple exercises.

Scaling the amount of feedback is harder: There is a strong dependency between the number and size of test cases and the run time of the grading process. There is also a dependency between the size of test cases and the readability of the feedback. Hence using more and larger test cases is not always helpful in means of better feedback. This can partially be solved by providing filtering information, so that potentially uninteresting parts of the test output (which can e.g. be console output from the program under test or program traces produced by JACK) are removed. Careful design of test cases can also help to reduce the amount of unnecessary feedback. This in turn possibly increases the setup time for an exercise. Another way of tackling this problem is postprecessing of test output, e.g. by highlighting relevant sections. However, this adds an extra amount of run time and thus decreases performance. In summary, there is a trade-off and possibly also a scaling problem related to the amount of feedback, the number of test cases and the performance of the system.

3 Conclusions

This paper reported on four additional years of experience with JACK and the underlying architecture. It could be shown that JACK is scalable in various dimensions. One way of measuring sustainability is asking the students for their attitude towards the system. When JACK was first introduced, students were sceptical and some actively tried to avoid the system. But very soon the attitude turned and students engaged in supporting JACK, e.g. by becoming tutors. To get a more precise picture, students for the introductory course were invited to take part in a survey at the end of each winter term and answer questions regarding their attitude towards JACK. In almost every survey more than 60% of the participants expressed a positive or very positive attitude towards the respective system. This can be seen as a clear indicator that the use of JACK has a solid and sustainable support from the students using it.

References

1. Ihantola, P., Ahoniemi, T., Karavirta, V., Seppälä, O.: Review of recent systems for automatic assessment of programming assignments. In: Proceedings of the 10th Koli Calling International Conference on Computing Education Research, Koli Calling 2010, pp. 86–93. ACM, New York (2010)
2. Striewe, M., Balz, M., Goedicke, M.: A flexible and modular software architecture for computer aided assessments and automated marking. In: Proceedings of the First International Conference on Computer Supported Eductation (CSEDU), INSTICC, Lisboa, Portugal, March 23 - 26, vol. 2, pp. 54–61 (2009)

Evaluating Relevance of Educational Resources of Social and Semantic Web

Davide Taibi[1], Giovanni Fulantelli[1], Stefan Dietze[2], and Besnik Fetahu[2]

[1] Istituto per le Tecnologie Didattiche, Consiglio Nazionale delle Ricerche, Palermo, Italy
{davide.taibi,giovanni.fulantelli}@itd.cnr.it
[2] L3S Research Center, Hannover, Germany
{dietze,fetahu}@l3s.de

Abstract. The social web paradigm has modified the way people behave on the Web. Amongst the many consequences of this change the amount of online resources directly produced and shared by users has increased considerably. In this scenario the importance of methods to evaluate the educational relevance of the resources raises up. In this poster we propose an approach based on recent advancements of Linked Open Data.

Keywords: Linked Open Data, OER, Educational relevance of resource.

1 Evaluating Educational Relevance: A Linked Data Approach

In the last few years the development of the Social Web has changed the way Open Educational Resource repositories are considered, from static collectors of resources to dynamic and social environments able to support social activities such as sharing of knowledge, commenting and voting. At the same time, social media environments, such as YouTube, Flickr, Slideshare, have included specific sections related to education, thus becoming a meaningful source of educational resources. In this scenario the problem of promoting an automated approach to analyze and evaluate the huge amount of resources suitable for learning raised up. Duval and Ochoa have studied relevance ranking metrics for Learning Objects (LOs), proposing the LearnRank function [1] based on Pagerank (the well-known algorithm used by search engines to evaluate the relevance of web pages) and a model to transform the relevance concept into a numerical value [2]. Both authors consider relevance of LOs as a multi-dimensional issue, where the learning context plays a key role. While these studies focus on the relevance of LOs, the recent evolution of Social Web provides further potentials to infer relevant clues about the educational relevance of resources on social media environments by exploiting social activities such as rating, liking and commenting; furthermore, the concept of learning context is affected as well; finally, the increasing publication of datasets in the background of the Semantic Web requires a review of the previous models to evaluate the educational relevance of resources.

In the framework of the LinkedUp project (http://linkedup-project.eu/), several *data curation* activities have being carried out, aimed at assessing, cataloging, annotating and exposing all sorts of Web data of educational relevance. Both explicitly

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 637–638, 2013.

educational datasets (such as OpenLearn and the mEducator Educational Resources) as well as implicitly educationally relevant datasets (e.g. BBC Programmes and Europeana) have been selected in order to create the Linked Education Catalog. Due to the heterogeneity of the resources selected, a *classification* of the available datasets, indicating their *main purpose*, *nature* and *educational relevance* is required. Amongst the data curation activities [3][4], in [4] the authors presented an approach, based on a set of data processing techniques to enable resource classification by representing relationships between resources as a network and using cluster algorithms. In this poster the proposed approach has been extended, in order to exploit the properties of the network, such as degree, centrality measures, to provide a relevance measure for learning resources. Moreover, votes, comments, number of visualizations, and other social activities related to each resource can also be used to weight the relationships between resources, thus yielding relevance index that takes into account both social and semantic aspects. Figure 1 shows a part of the resulting network including example resources from BBC, Linked Universities and Europeana datasets.

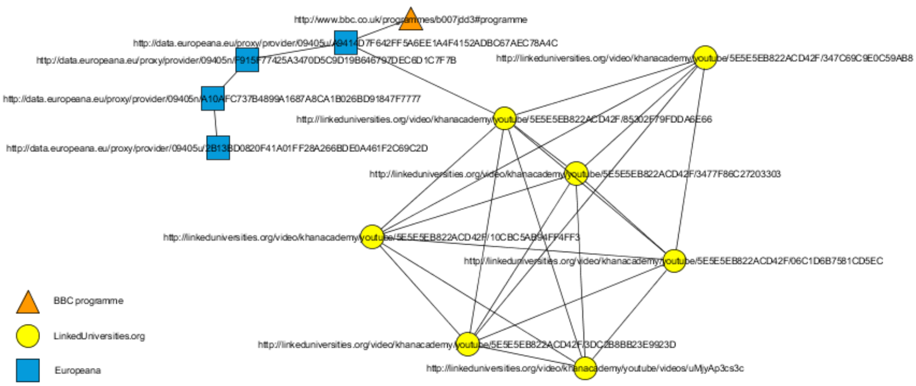


Fig. 1. A network of learning resources

References

1. Duval, E.: Policy and Innovation in Education – Quality Criteria, chapter LearnRank: the Real Quality Measure for Learning Materials. European Schoolnet, 457–463 (2005)
2. Ochoa, X.: Learnometrics: metrics for learning objects. In: Proc. of the 1st Int. Conf. on Learning Analytics and Knowledge (LAK 2011), pp. 1–8. ACM, New York (2011)
3. d'Aquin, M., Adamou, A., Dietze, S.: Assessing the Educational Linked Data Landscape. In: ACM Web Science (WebSci 2013), Paris, France (May 2013)
4. Taibi, D., Fetahu, B., Dietze, S.: Towards Integration of Web Data into a coherent Educational Data Graph. In: ACM Proceedings of Linked Learning 2013: the 3rd International Workshop on eLearning Approaches for the Linked Data Age (2013)

Designing Computational Systems for Serendipity in Learning

Maria Taramigkou[1], Fotis Paraskevopoulos[1], Efthimios Bothos[1], Dimitris Apostolou[2], and Gregoris Mentzas[1]

[1] National Technical University of Athens
{martar,fotisp,mpthim,gmentzas}@mail.ntua.gr
[2] University of Piraeus
dapost@unipi.gr

Serendipity, the process of making fortunate discoveries for which someone was not looking for, can play a crucial role in leveraging creativity in learning [1]. Serendipity enables creative connections to develop while it can have a role in revealing hidden connections or "hidden analogies", especially in a social context such as in most learning processes [2]. The results of a chance encounter can result in new ideas relevant to the learner's previous knowledge [3]. In previous work, we have reviewed the related literature and identified five enabling factors for serendipity [4]: *diversity*, *unexpectedness* or *novelty*, *personalization*, *visualization*, and *social interaction*. The aim of this paper is to construct an architectural framework based on the aforementioned factors that can be used as a guide in the development of information seeking systems aiming to leverage serendipity in learning.

We consider the information seeking process as an iterative process in which the learner seeks inspiring, serendipitous information within layers of the available resource space comprising enterprise and Web sources that are possible anchors to serendipitous findings. We have separated the sources of information in the following categories: Inspirational Artifacts (0) consisting of pools of inspirational resources such as idea repositories that may be available to learners. Social Networks (1) consists of popular services where the "who follows whom" paradigm exists such as LinkedIn. Search Engines (2) can be utilised for augmenting data retrieved from information streams. Sensors (7) in the form of e.g., browser plugins can gather information from the users activity whilst using their browser (3) of preference. These sensors will track the URLs the learner visits thus giving us access to the content, which in turn can be used for processing down the processing pipeline. Tools can be used to gather information from internal sources such as proxies (8), logs (9) or crawling (10). Crawling caters more for historical data rather than real-time streams of information. Finally, we may have access to other artefacts that can inspire learners in the form of conceptual designs, best practices, ideas from similar domains, etc.

Having defined the information sources, the next step is to determine what type of information is accessible and build modules which can handle the data processing and storage. For example, when gathering information from Social Networks we need to handle "live" streams of information and process them in real-time, possibly employing event possessing capabilities, whereas Search Engines are processed on demand where the search query is formulated in real-time. Furthermore, sensors gather information every second and then store it at specific intervals.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 639–640, 2013.

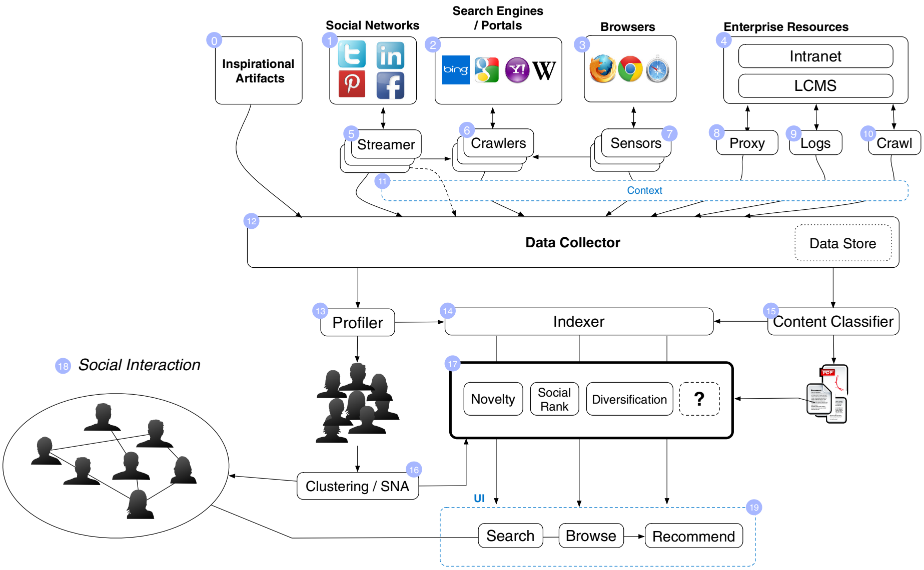


Fig. 1. Architectural Framework for Supporting Serendipity in Learning

The Data Collector (12) module takes care of providing a unified interface so that streamer, crawl and sensor modules store information using the same mechanisms. Once the information is stored, some pre-processing is needed by means of the Profiler (13), which will model the user, to determine her areas of expertise, her interests, etc.; the Content Classifier (15), which will pre-process the information to determine common traits between content regardless of information sources; and the Indexer (14), which will pre-process the results of common queries when searching for profiles or content stored in the component. The architecture also includes a module for extracting, modelling and storing "Social Interaction" (18) data, e.g., clustering profiles in order to create groups (16). The "Logic" (17) set of modules is considered to be the "intelligent" part of the architecture that contain the algorithms, which perform diversification or boost novelty of content and provides the results for inspirational search, browsing as well as recommendations. These results are accessed via the UI (19) which will provide the user with alternate interfaces to assist in the creativity process for the goal at hand.

References

1. Foster, A., Ford, N.: Serendipity And Information Seeking: An Empirical Study. Journal of Documentation 59(3), 321–340 (2003)
2. Cory, K.A.: Discovering hidden analogies in an online humanities database. Computers and the Humanities 31(1), 1–12 (1997)
3. Boden, M.A. (ed.): Dimensions of creativity. MIT Press (1996)
4. Taramigkou, M., Bothos, E., Apostolou, D., Mentzas, G.: Serendipity in Online Information Systems. In: Proceedings of the 2013 IEEE International Technology Management Conference & 19th ICE Conference, June 24-26. The Hague (2013)

Designing Scalable Informal Learning Solutions with Personas: A Pilot Study in the Healthcare Sector

Stefan Thalmann[1], Vanessa Borntrager[1], Tamsin Treasure-Jones[2], John Sandars[2], Ronald Maier[1], Kathrin Widmann[1], and Micky Kerr[2]

[1] University of Innsbruck School of Management, Innsbruck, Austria
[2] Leeds Institute of Medical Education, University of Leeds, Leeds, UK

1 Motivation

Driven by ever shorter cycles of innovation, organizations and individuals nowadays have to acquire, understand and apply new knowledge in shorter periods of time [1]. Much of this rapid learning appears to be achieved by workers learning on the job and from colleagues – informal learning rather than learning from traditional, curriculum-based training [2]. Mobile technology could potentially provide support to this informal learning as it can provide scalable and flexible learning tools that can be used at any time across a variety of locations: at home, on different work sites, during travel [3]. However, designing learning technology that can support such unstructured, creative and expertise-driven informal learning is challenging, especially as there are likely to be great variations across employees in terms of their perceptions, experiences and expectations regarding technology [4]. These expectations and experiences may also differ from those of the developers and designers. Yet it is the match between user requirements and functionalities that lies at the heart of a successful learning technology. In Learning Layers we are exploring how creating and using Personas may help to design scalable technology for supporting informal learning in healthcare.

2 Background

We have been working with medical practices to understand their informal learning context. This area is a good fit for our interest in informal learning since professional practice in the healthcare sector is characterized by dealing with complex situations that require the skilful integration of knowledge [2]. Generalized information is often of little practical value for effectively dealing with the 'messy' situations in daily practice [5]. In such circumstances, greater value is found in the tacit knowledge, either from individual experience or from the experience of others. Furthermore for sharing these experiences these independent organisations have to agree to adopt cross-organizational learning solutions. That in turn requires designers to understand what is acceptable for users as well as what is technically possible. To obtain this rich understanding of current informal learning practice we ran 3 focus groups with 3 UK medical practices, involving 23 participants. Based on an analysis of the transcripts Personas were created: http://learning-layers.eu/layers-personas-for-the/

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 641–642, 2013.

3 Results

In analyzing the data and developing the Personas, we observed an increased demand for communication and knowledge acquisition concerning informal learning. But it also becomes clear that employees are currently not sufficiently supported in their learning and knowledge acquisition. In particular more senior staff (e.g. Personas: Natasha and George), seem overextended with the growing communication demand caused by their colleagues who strive for more knowledge. Hence, these Personas are overwhelmed with redundant knowledge sharing activities, yet they are also key actors for informal learning activities in their networks and supporting their knowledge creation and sharing work with ICT could have a big impact across the organization. Building on this understanding we are collaboratively developing solutions for scalable learning technologies to support these Personas. Initial design ideas include:

A tool to facilitate easy content production and sharing capabilities, including the ability to track maturity of the content and share with trusted networks.

A tool supporting individuals in keeping track of their learning needs as they arise during work. Providing a space where they can reflect on and respond to these learning prompts and choose to share this learning with others. This requires providing a tool that can help them to enrich, revise, structure, assemble, search and share these contents collaboratively. The core idea is to integrate learning and sense making into the work environment.

Acknowledgments. The research leading to the presented results was partially funded by the EC under the 7th Framework Programme through the project LEARNING LAYERS.

References

1. Wetherill, M., Rezgui, Y., Lima, C., Zarli, A.: Knowledge Management for the Construction Industry: The E-Cognos Project. Journal of Information Technology in Construction (7), 183–196 (2002)
2. Eraut, M.: Informal learning in the workplace. Studies in Continuing Education 26(2), 247–273 (2004)
3. Wang, M., Shen, R.: Message Design for Mobile Learning: Learning Theories, Human Cognition and Design Principles. British Journal of Educational Technology 43(4), 561–575 (2011)
4. Maier, R., Thalmann, S.: Using Personas for Designing Knowledge and Learning Services: Results of an Ethnographically Informed Study. International Journal of Technology Enhanced Learning 2(1/2), 58–74 (2010)
5. Schön, D.A.: The reflective practitioner: How professionals think in action, vol. 5126, basic books (1983)

Enabling Educators to Design Serious Games – A Serious Game Logic and Structure Modeling Language*

Niroshan Thillainathan, Holger Hoffmann, Eike M. Hirdes, and Jan Marco Leimeister

Kassel University, Information Systems, Kassel, Germany
{thillainathan,hhoffman,hirdes,leimeister}@uni-kassel.de

1 Introduction

Serious games are applications combining educational content with gameplay by integrating learning objectives into a game-like environment to keep up the player's motivation to continue playing, and hence learning. This characteristic is highly sought after in educational contexts, making serious games a big asset for didactics [1]. Offering new learning contents through a game not only induces higher motivation, employing serious games can also yield higher learning success than presenting material in a classical, non-computer based, way [2]. Only few people having the proper didactical background to tailor the learning objectives to the students' need also have the programming knowledge and game design skills allowing them to develop didactically and technically sound serious games [3, 4]. In this paper, we argue for an approach to enable didactical experts, i.e. educators, to develop serious games adapted to their own learning content. To address this problem we develop a tool allowing educators to visually design their serious games, which is based on model driven development techniques that allow the generation of software from visual models. We describe the first step towards this tool, the development of the underlying domain specific modeling language (DSML).

2 GLiSMo: Serious Game Logic and Structure Modeling Language

To identify requirements for the development of our domain specific modeling language GLiSMo an in-depth literature review was conducted. As the result of our review, we have identified 18 publications mentioning 23 influencing factors for serious games, from which we have derived 24 requirements. GLiSMo is a modeling language, which allows designing the structure as well as the logic of a serious game.

Serious game structure modeling describes how a serious game and its game world are built. Here, the *Serious Game Root* element, present once for each game modeled, plays a special role, as it is the point of origin for all other elements. Serious game

* Parts of this research were funded by the German Federal Ministry of Education and Research and the European Social Fund in the project BlendedContENT (www.blendedcontent.de), FKZ 01PF08022A.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 643–644, 2013.

root elements can have one or more acts, which have similarities to levels in common games and are divided by one or more *scenes*. A scene describes a specific spot or place within the game world. This can for example be a room within a building. The element *Object* represents all kind of objects within the game world, i.e. doors, boxes, trees. The *Character* element represents the player itself and non-player characters (NPC). An *inventory* element is used for storing objects taken by the player. The element *feedback manager* manages the textual feedback given to the player. The *Reward Manager* stands for the scoring and rewarding within a serious game. The *Audio Manager and Video Manager* allow the playback of audio and video during the game. The *GUI Manager* is used to display buttons, textboxes, scores or the game menu.

Serious game logic modeling characterizes the behavior in terms of how does the game react on specific actions performed by the player, or events occurring during game play. It also describes the assessment of player actions and the game adaption according to the adaption results. To define the boundaries of a logic model, we denote the *initial-state* by a black circle; *final-states* are represented by an encircled black circle. The element *action* enables the player to interact with the game world. This has different types, e.g. *take object, use object,* representing different interaction methods in the game world. The element *task* represents different tasks, i.e. multiple-choice questions. The associated *assessment* element evaluates the result and initiates processes to give feedback and achievements to the player. To manage the information and control flow, we use so called *streams* to connect the elements. To make a progress along the control flow the player must execute the described actions. To send *events and messages* between elements a dotted arrow will be used. Furthermore *fork and join* as well as *branch* elements are used to manage parallel and branched flows.

3 Conclusion and Future Work

Supporting educators design serious games through model driven development techniques is a promising approach. In this paper we proposed our first steps towards a DSML for serious games. Future work includes the extension of the modeling language with more elements supporting different tasks and assessment methods. Furthermore GLiSMo must be evaluated with educators to determine its usability. In parallel a visual design tool for GLiSMo will be implemented.

References

1. Brennecke, A., Schumann, H.: A General Framework for Digital Game-Based Training Systems. In: IADIS International Conference Game and Entertainment Technologies, Algarve, Portugal (2009)
2. Pivec, M.: Editorial: Play and learn: potentials of gamebased learning. British Journal of Educational Technology 38(3), 387–393 (2007)
3. Chandler, H.M.: The Game Production Handbook, 2nd edn. Jones & Bartlett Publishers, Ontario (2010)
4. Torrente, J., et al.: <e-Adventure>: Introducing educational games in the learning process. in Education Engineering, EDUCON (2010)

Re-engineering the Uptake of ICT in Schools

Frans Van Assche[1], Bernd Simon[2], Michael Aram[2], Jean-Noël Colin[3], Hoang Minh Tien[3], Dai Griffiths[4], Kris Popat[4], Luis Anido-Rifón[5], Manuel Caeiro-Rodríguez[5], Juan Santos-Gago[5], Will Ellis[6], and Joris Klerkx[1]

[1] University of Leuven, Computer Science Department, Belgium
{frans.vanassche,joris.klerkx}@cs.kuleuven.be
[2] Knowledge Markets, Austria
{bernd.simon,michael.aram}@km.co.at
[3] University of Namur, Belgium
{jean-noel.colin,minh-tien.hoang}@unamur.be
[4] University of Bolton, United Kingdom
{dai.griffiths.1,kjpopat}@gmail.com
[5] University of Vigo, Spain
{lanido,mcaeiro,jsgago}@det.uvigo.es
[6] European Schoolnet, Belgium
will.ellis@eun.org

Abstract. While many innovations in Technology Enhanced Learning (TEL) have emerged over the last two decades, the uptake of these innovations has not always been very successful, particularly in schools. The transition from proof of concept to integration into learning activities has been recognized as a bottleneck for quite some time. This major problem, which is affecting many TEL stakeholders, is the focus of the four year iTEC project that is developing a comprehensive approach.

Keywords: Systemic change, uptake of ICT, schools.

Following the introduction of the World Wide Web, resulting in a boost in innovations in the field of Technology Enhanced Learning (TEL), many (more) barriers to the uptake of ICT in schools have been identified. The importance of this problem can not be overestimated given the wide range of stakeholders that are affected by it and the ensuing lack of adoption of ICT in schools. Therefore, a comprehensive approach is required in order to change the way key players, including teachers, learners, ICT co-ordinators, and policy makers adopt emerging technologies. Such an approach, developed in the iTEC[1] project, is built on the following five principles.

First, *addressing the mainstreaming gap* given that the effectiveness of the mainstreaming is often the most determining factor in changing the practice for capitalizing on what ICT can offer. Second, *aiming for systemic change* by improving the mainstreaming *approach* rather than to provide a few successful mainstreaming show

[1] iTEC stands for innovative Technologies in an Engaging Classroom. The iTEC project is partially funded by the European Commission under the seventh Framework Programme.

D. Hernández-Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 645–646, 2013.

cases of hyped technology that might be obsolete in a few years. Third, *connecting with current practice of learners, teachers, head teachers, and policy makers* as designs for the future classroom that are simply too unconnected with current practice, fail to engage teachers and cannot be mainstreamed because they are divorced from educational policy making in the real world. Fourth, *building on the engaging potential of emerging technologies for learners and teachers* and fifth, *innovation with beneficial impact* viz. innovation must be a change that creates positive value and is better (more effective, more efficient) than its predecessor.

Based on these principles, iTEC is developing the following innovative artefacts:

- *Ready-Made Scenarios.* iTEC provides a set of scenarios (i.e. de-contextualised structured narrative learning designs) that are aimed to help teachers to go beyond their usual classroom activities and to explore emerging technologies. These scenarios are adapted by teachers to their own local context.
- *Ready-Made Learning Activities.* Learning stories consist of learning activities and are further elaborations of scenarios as concrete instantiations. Learning activities make the resource (material, people, events) requirements more concrete. By providing different levels of abstraction, teachers and learners can choose the appropriate level for their purpose.
- *A Future Classroom Scenario Method.* As iTEC is aiming for systemic change, it also provides a method with procedures and techniques for developing such scenarios. An important part of this toolkit is the iTEC innovation matrix that allows teachers, head-teachers, ICT co-ordinators, and MoE to assess where they are with respect to four innovation dimensions, and develop scenarios that facilitate taking the next step.
- *The Learning Activity Design Method*, that guides teachers in how to find and use an archive of Learning Stories and Learning Activities which are derived from iTEC scenarios. It is much more focused on enabling the adoption of advanced pedagogical approaches by teachers, supported by appropriate technologies and other resources. The Learning Activity Design Toolkit is used by individual teachers and collaborative communities.
- *A Collection of Widgets.* The iTEC widgets are small ready-to-use educational apps that can be deployed in different so-called ‘shells’. A shell is a container technology for widgets.
- *A Number of Technical Artefacts, Including Services and Specifications.* These artefacts, include automated help in finding feasible scenarios, automated support for localisation, finding more easily other types of resources, download and play applications, plug and play authentication and authorisation, support in establishing new collaborations, and last but not least the iTEC Educational Cloud (IEC).

All these iTEC artefacts have one characteristic in common: facilitating the uptake of ICT in schools. However, the benefits are not restricted to that and the iTEC technical artefacts will also be beneficial to other stakeholders including technology providers, standardization bodies, and researchers.

Author Index