Christoph Rensing
Sara de Freitas
Tobias Ley
Pedro J. Muñoz-Merino (Eds.)

LNCS 8719

# Open Learning and Teaching in Educational Communities

9th European Conference on Technology Enhanced Learning, EC-TEL 2014
Graz, Austria, September 16–19, 2014, Proceedings

Springer

# Lecture Notes in Computer Science 8719

*Commenced Publication in 1973*
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen


## Editorial Board

David Hutchison
*Lancaster University, UK*
Takeo Kanade
*Carnegie Mellon University, Pittsburgh, PA, USA*
Josef Kittler
*University of Surrey, Guildford, UK*
Jon M. Kleinberg
*Cornell University, Ithaca, NY, USA*
Alfred Kobsa
*University of California, Irvine, CA, USA*
Friedemann Mattern
*ETH Zurich, Switzerland*
John C. Mitchell
*Stanford University, CA, USA*
Moni Naor
*Weizmann Institute of Science, Rehovot, Israel*
Oscar Nierstrasz
*University of Bern, Switzerland*
C. Pandu Rangan
*Indian Institute of Technology, Madras, India*
Bernhard Steffen
*TU Dortmund University, Germany*
Demetri Terzopoulos
*University of California, Los Angeles, CA, USA*
Doug Tygar
*University of California, Berkeley, CA, USA*
Gerhard Weikum
*Max Planck Institute for Informatics, Saarbruecken, Germany*

Christoph Rensing Sara de Freitas
Tobias Ley Pedro J. Muñoz-Merino (Eds.)

# Open Learning and Teaching in Educational Communities

9th European Conference
on Technology Enhanced Learning, EC-TEL 2014
Graz, Austria, September 16-19, 2014
Proceedings

Springer

Volume Editors

Christoph Rensing
Technische Universität Darmstadt
Multimedia Communications Lab.
Rundeturmstr. 10, 64283 Darmstadt, Germany
E-mail: christoph.rensing@kom.tu-darmstadt.de

Sara de Freitas
Murdoch University
Vice Chancellery, 90 South Street, Murdoch, WA 6150, Australia
E-mail: s.defreitas@murdoch.edu.au

Tobias Ley
Tallinn University
Center for Educational Technology
Narva Maantee 29, 10120 Tallinn, Estonia
E-mail: tley@tlu.ee

Pedro J. Muñoz-Merino
Universidad Carlos III de Madrid
Department of Telematic Engineering
Avenida Universidad 30, 28911 Leganés (Madrid), Spain
E-mail: pedmume@it.uc3m.es

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-319-11199-5 e-ISBN 978-3-319-11200-8
DOI 10.1007/978-3-319-11200-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014947383

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

© Springer International Publishing Switzerland 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in ist current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

*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

The ninth edition of the European Conference on Technology Enhanced Learning (EC-TEL) was held in Graz (Austria) during September 16-20, 2014. This volume collects all peer-reviewed contributions that were included in the exciting program of this year's conference.

EC-TEL is growing each year and is becoming the major interdisciplinary venue for the community of technology-enhanced learning researchers in Europe and worldwide. Since 2006, EC-TEL has provided a reference point 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), Saarbrücken (Germany, 2012) and Paphos (Cyprus, 2013).

The theme of EC-TEL 2014 was "Open Learning and Teaching in Educational Communities." This is to recognize that educational and learning arrangements in our modern society are increasingly characterized by openness in terms of participation, community building as well as highlighting the importance of tools, devices, and resources used to support open access and social learning. As part of a greater drive toward social learning, we note shifts between formal, non-formal, and informal group learning.

Today, learners are coming together in different social settings and communities such as in groups in schools, teams in a company, and in informal communities of learners. They share and co-create educational resources and interact with an ever increasing number of open educational resources (OER). Learning communities differ in dimension from local ones, through local groups to widely distributed global communities, such as witnessed by the emerging MOOCs paradigm. The learners use multiple computational tools and devices, often at the same time, e.g., tablets and smartphones for individual learning, interactive tables and electronic whiteboards for collaboration, or mobile technologies for taking learning outside the classroom. Often, access to the open educational arrangements take place in a mobile and ubiquitous fashion.

To understand how learning in such open educational communities happens and how learning and teaching technology can be supported in this context was the guiding question in this year's edition. Drawing on the core TEL disciplines of computer science, education, psychology, cognitive science and social science, research contributions presented in EC-TEL 2014 addressed topics such as informal learning, self-regulated and self-directed learning, reflective learning, inquiry-based learning, communities of learners and communities of practice, learning design, learning analytics, personalization and adaptation, social media, computer-supported collaborative learning, massive open online courses (MOOCs), and schools and universities of the future.

Moreover, EC-TEL 2014 was collocated with the 14th International Conference on Knowledge Technologies and Data-Driven Business (i-KNOW 2014) and the programs were tightly integrated with shared plenary and keynote sessions. This introduced the EC-TEL community to emerging topics such as data analytics, which is also shaping the learning scene as education becomes more open and accessible.

This 2014 edition was extremely competitive, given the high number of submissions generated yet again. Out of the over 200 initial abstract submissions, a total of 165 valid paper submissions were received. Of these, 113 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 meta-reviews were elicited if necessary. Finally, 27 submissions were selected as full papers (resulting in an acceptance rate for full papers of 23.9%). Additionally, 18 papers were presented as short papers, 16 as demonstrations and 26 as posters. Table 1 shows the detailed statistics.

**Table 1.** Acceptance rate in different submission categories

| | | Accepted as | | | |
|---|---|---|---|---|---|
| Submitted as | | Full Paper | Short Paper | Poster Paper | Demo Paper |
| Full Paper | 113 | 27 | 15 | 11 | 6 |
| Short Paper | 28 | | 3 | 8 | 1 |
| Poster Paper | 9 | | | 5 | |
| Demo Paper | 15 | | | 2 | 9 |
| Sum | 165 | 27 | 18 | 26 | 16 |

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 high number of submissions and still keep all deadlines as originally planned.

Keynote presentations completed this competitive scientific program, Etienne Wenger-Trayner, educational theorist and practitioner, best known for his seminal work on communities of practice and social learning theory, gave a talk entitled "Open Learning and Teaching: Perspectives from the Social Learning Theory." In addition there were two keynotes shared with the i-KNOW conference. Bernardo A. Huberman, Senior HP Fellow and Director of the Social Computing Lab at Hewlett-Packard Laboratories, presented his work on social attention. Viktor Mayer Schönberger, Chair of Internet Governance and Regulation at the Oxford Internet Institute, gave a talk about big data and how it promises to fundamentally change how we make sense of the world.

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 Shootout" in which the demonstrations were shown "in action" giving the audience the possibility to vote for the best demo. Demonstrations were also interactively exhibited during the conference and posters during the welcome reception, sparking discussions between different but complementary groups, bringing the community closer to solving the many 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. In all, 11 workshops and tutorials were selected from the proposals and were organized. Some of them continue a row of established workshops, like the workshops on motivational and affective aspects in TEL or on awareness and reflection in TEL. Others consider new topics like MOOCs or smart city learning. A doctoral consortium was organized concurrently with the workshops. It provided an opportunity for PhD students to discuss their work with experienced TEL researchers.

We would like to thank the many contributors without whom the conference would not have been possible. These include foremost the authors, the PC members and reviewers, and the conference chairs, who all contributed to the program. We would also like to thank an enthusiastic and dedicated local organization team in Graz who made EC-TEL a smooth and memorable experience. The conference was also partially supported by the European Association of Technology-Enhanced Learning (EATEL), the University of Graz, Springer, EasyChair, and the *IEEE Transactions on Learning Technologies.*

September 2014

Christoph Rensing
Sara de Freitas
Tobias Ley
Pedro J. Muñoz-Merino

# Organization

## Executive Committee

## General Chair

| | |
|---|---|
| Tobias Ley | Tallinn University, Estonia |

### Program Chairs

| | |
|---|---|
| Sara de Freitas | Murdoch University, Australia |
| Christoph Rensing | Technische Universität Darmstadt, Germany |

### Industry Chairs

| | |
|---|---|
| Nils Faltin | imc information multimedia communication AG, Germany |
| Jon Arambarri | Virtualware, Spain |

### Workshop Chair

| | |
|---|---|
| Viktoria Pammer-Schindler | Graz University of Technology, Austria |

### Poster and Demonstration Chair

| | |
|---|---|
| Pedro J. Muñoz-Merino | Universidad Carlos III de Madrid, Spain |

### Doctoral Consortium Chairs

| | |
|---|---|
| Katherine Maillet | Telecom SudParis, France |
| Vanessa Camilleri | University of Malta, Malta |

### Dissemination Chair

| | |
|---|---|
| Hendrik Drachsler | Open University of the Netherlands, The Netherlands |

### Local Organization Chair

| | |
|---|---|
| Stefanie Lindstaedt | Graz University of Technology, Austria |

## Program Committee

Marie-Helene Abel
Achilleas Achilleos
Mohammad Al-Smadi
Carlos Alario-Hoyos
Liaqat Ali
Mario Allegra
Luis Anido Rifon
Inmaculada Arnedillo-Sanchez
Juan I. Asensio-Perez
Francesco Bellotti
Jason Bernard
Martina Bientzle
Katrin Borcea-Pfitzmann
Yolaine Bourda
Andreas Breiter
Francis Brouns
Peter Brusilovsky
Manuel Caeiro
Vanessa Camilleri
Lorenzo Cantoni
Manuel Castro
Mohamed Amine Chatti
John Cook Audrey Cooke
Raquel M. Crespo García
Alexandra Cristea
Mihai Dascalu
Silvia De Ascaniis
Paul De Bra
Sara de Freitas
Carlos Delgado Kloos
Stavros Demetriadis
Christian Depover
Michael Derntl
Philippe Dessus
Darina Dicheva
Stefan Dietze
Yannis Dimitriadis
Vania Dimitrova
Juan Manuel Dodero
Silvester Draaijer
Hendrik Drachsler
Jon Dron
Benedict Du Boulay
Erik Duval
Martin Ebner
Nils Faltin
Baltasar Fernandez-Manjon
Alejandro Fernández
Carmen Fernández-Panadero
Beatriz Florian-Gaviria
Katharina Freitag
Muriel Garreta
Dragan Gasevic
Manuel Gentile
Sebastien George
Adam Giemza
Denis Gillet
Fabrizio Giorgini
Carlo Giovannella
Christian Glahn
Sabine Graf
Monique Grandbastien
Andrina Granić
David Griffiths
Begona Gros
Christian Gütl
Joerg Haake
Iassen Halatchliyski
Andreas Harrer
Davinia Hernández-Leo
Knut Hinkelmann
Ulrich Hoppe
I-Han Hsiao
Christoph Igel
Hazra Imran
Seiji Isotani
Isa Jahnke
Marco Kalz
Georgia Kapitsaki
Nuri Kara
Nikos Karacapilidis
Michael Kickmeier-Rust
Andrea Kienle
Barbara Kieslinger
Ralf Klamma
Joris Klerkx

Tomaž Klobučar
Johannes Konert
Vitomir Kovanovic
Milos Kravcik
Barbara Kump
Mart Laanpere
Lydia Lau
Elise Lavoué
Effie Lai-Chong Law
Danielle H. Lee
Derick Leony
Tobias Ley
Stefanie Lindstaedt
Andreas Lingnau
Martin Llamas-Nistal
Ulrike Lucke
Moushir M. El-Bishouty
George Magoulas
Katherine Maillet
Nils Malzahn
M. Antonia Martínez-Carreras
Javier Melero
Martin Memmel
Agathe Merceron
Riichiro Mizoguchi
Yishay Mor
Pablo Moreno-Ger
Pedro J. Muñoz-Merino
Mario Muñoz-Organero
Kati Mäkitalo-Siegl
Rob Nadolski
Ambjorn Naeve
Jelena Nakić
Roger Nkambou
Viktoria Pammer
Pantelis Papadopoulos
Abelardo Pardo
Kai Pata
Zinayida Petrushyna
Niels Pinkwart
Michael Prilla
Mar Pérez-Sanagustín
Eric Ras
Christoph Rensing
Verónica Rivera-Pelayo
Luz Stella Robles Pedrozo
Miguel Rodriguez-Artacho
Demetrios Sampson
Patricia Santos
Andreas Schmidt
Judith Schoonenboom
Ulrik Schroeder
Till Schümmer
Mike Sharples
Bernd Simon
Nina Simon
Sergey Sosnovsky
Marcus Specht
Slavi Stoyanov
Kairit Tammets
Stefano Tardini
Deborah Tatar
Pierre Tchounikine
Stefaan Ternier
Vladimir Tomberg
Christoph Trattner
Stefan Trausan-Matu
Aimilia Tzanavari
Katrien Verbert
Terje Väljataga
Jo Wake
Armin Weinberger
Wim Westera
Fridolin Wild
Martin Wolpers
Sabrina Ziebarth

## Additional Reviewers

Antonio Balderas
Mohammad Hassan Falakmasir
Vera Gehlen-Baum
Franka Gruenewald
Julio Guerra
Raluca Judele
Micky Ker
Styliani Kleanthous Loizou
Dominik Kowald
Jeff Kurcz
Emanuel Lacic
Sven Manske
Iza Marfisi-Schottman
Jochen Rick
Stylianos Sergis
Júlia M. C. Silva
Richard Tortorella
Sara Vannini
Panagiotis Zervas

# Table of Contents

## Full Papers

## Short Papers

## Demo Papers

## Poster Papers

# Eliciting Requirements for Learning Design Tools

## A Semio-Participatory Approach


Alessandro Arpetti[1,2], M. Cecília C. Baranauskas[1], and Tommaso Leo[2]

[1] Institute of Computing, University of Campinas, UNICAMP, Campinas, Brazil
{alessandro.arpetti,c.baranauskas}@ic.unicamp.br
[2] Department of Engineering Information, Marche Polytechnic University, Ancona, Italy
{a.arpetti,tommaso.leo}@univpm.it



**Abstract.** The need to properly design educational intervention, representing explicitly what students and teachers are planned to do, has been acknowledged in literature. Thus, Learning Design (LD) tools, if made accessible and usable by teachers, can bring significant benefits potentially improving results of educational practices. Although effort has been made in developing systems to support the learning design process, literature has shown they have not yet reached a sufficient spread among teachers. This paper investigated the subject by conducting semio-participatory practices with a group of teachers at a distance, to understand the meaning they make to issues regarding learning design practices and representations, aiming at eliciting user requirements for a prospective LD tool.



**Keywords:** Learning Design, Requirements Analysis, Semio-Participatory Practices.


## 1 Introduction

In recent decades, significant socio-cultural changes and the rapid evolution of information and communication technologies have significantly modified the educational scenario, introducing greater complexity and numerous challenges to ensure the effectiveness of education. Regarding the use of technology in education, it has been acknowledged the need to properly design educational interventions, representing explicitly what students and teachers are planned to do (Beetham, 2007). In this scenario, Learning Design (LD), i.e., the design of educational actions, is a key factor that, if made accessible and usable by teachers, can bring significant benefits potentially improving results of educational practices (Lockyer, Bennett, Agostinho, & Harper, 2009).

The term "Learning Design" began to appear in the late 90's, in studies related to the Instructional Design field (Persico et al., 2013), although there is not yet a shared vocabulary within the research community. As highlighted by Dobozy (2011), the field itself is called "learning design" (Dalziel, 2006), "instructional design" (Chu & Kennedy, 2011), "curriculum design" (Ferrell, 2011), "educational design" (Goodyear


C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 1–14, 2014.


© Springer International Publishing Switzerland 2014

& Ellis, 2007), “design for learning” (Beetham & Sharpe, 2013) and “design-based learning” (Wijnen, 2000). Another relevant position defines the field as “pedagogical planning” (Earp & Pozzi, 2006; Gutierrez, Valigiani, Jamont, Collet, & Kloos, 2007). Agostinho (2006) provides a general definition for the process of LD as the representation of teaching and learning practices using a notational format. The aim of this practice is to create a plan of an educational intervention that can serve as model or template, adaptable by a teacher to suit his/her context and needs. Within a more technical point of view, Koper (2006) defines LD as the description of the teaching/learning process that occurs in a unit of learning (e.g., a course, a lesson or any other designed learning event). It represents the learning activities and the support activities that are performed by different persons (learners, teachers) in the context of a unit of learning. For this purpose, the IMS Learning Design specification aims to represent the LD in a semantic, formal and machine interpretable way. On the other hand, paying more attention to the sharing of experiences and professional growth of teachers, Conole (2013) defines LD as a methodology useful to guide teachers to make more informed decisions, through the elicitation of pedagogical and practical knowledge. This general definition is not restricted to units of learning, but includes the design of resources and individual learning activities right up to curriculum-level design. From this perspective, the main purpose of LD is to help make the design process more explicit and shareable. As a research area, LD includes both the understanding of the design process, as well as the development of LD resources, tools and activities.

Starting from IMS-LD specification (R. Koper, Olivier, & Anderson, 2003), many LD representations, software tools and design frameworks have been developed in the last years (Conole, 2013). Nevertheless, despite these efforts, no evidence has been presented yet regarding simplifying the design process or gaining a wider audience among teachers not specialized in LD or not proficient in the use of technology (Arpetti, Baranauskas & Leo, 2013a; Goodyear, 2005; Katsamani & Retalis, 2013; Oliver & Littlejohn, 2006).

This paper investigates the subject by conducting a study with teachers of Italian as second/foreign language, located in different countries, aimed at understanding the meaning they make to a prospective system intended to support their practices of LD. This work is part of the Ledita (Learning Design for Italian Language) research project (Arpetti, Baranauskas & Leo, 2013b) that aims at developing practical solutions and theoretical knowledge related to LD. The project is inspired by the Educational Based Research methodology (McKenney & Reeves, 2012) and is developed with the collaboration of a group of Italian language teachers. Following the first research phase, devoted to the analysis and exploration of the problem through a usability evaluation of the latest generation LD software tools and an investigation of teachers‟ design practices, this paper describes and discusses results of the predesign phase, which was intended to clarify the raised issues and specify user requirements for the development of a LD software tool.

In order to promote a better understanding of end user needs and develop solutions closer to their teaching realities, we adopted the Semio-Participatory approach, based on the assumption that “including the user in the design process is vital to make sure we are creating systems that make sense and that are part of the users‟ context of life" (Baranauskas, 2009). Inspired by Organizational Semiotics (Liu, 2000), the Semio-Participatory framework integrates the system design with social and participatory practices: the technical level of technology design (the software system) presupposes knowledge of formal (forms and rules) and informal (meanings, intentions, beliefs, responsibilities) social levels, understood by the analysis of signs carried by messages of participatory practices. Among the methods and artifacts proposed by these theories for problem analysis and requirements specification, we selected the Group Elicitation Method (GEM) (Boy, 1997) and the Problem Articulation Method (PAM) (Liu, 2000). The selection of these methods was motivated by their effectiveness in facilitating the communication among the participants on the problem clarification, definition and sharing of signs in their usefulness for the elicitation and specification of requirements. To facilitate the participation of teachers involved in the research project, we adapted these methods and artifacts for use in remote and asynchronous mode, through forums and shared editable documents in the project website. This paper describes the process and the results of the semio-participatory activities for the user requirement analysis of a LD software tool. The text is organized as follows: Section 2 illustrates the results of participatory requirement analysis activities; Section 3 presents the study findings; Section 4 discusses the results and Section 5 illustrates the conclusions of the study.

## 2 The Study

The Ledita project counts on about 90 participants, who are teachers of Italian as second/foreign language from 16 countries. All teachers had a Liberal Art education with a multidisciplinary background and most of them had a tertiary education and a multi-years teaching experience. They were asked about their relationship with technology and none of them said to have a negative one, whereas the majority declared to have a good or excellent relationship. The most commonly used technologies were computer, smartphone and tablet in private life, and computer and interactive whiteboard in professional life. With regard to their experience with software tools, all the participants used Internet; most of them had an e-mail address, an office software suite and some of them a graphics program and a video-editing tool.

In this study, interactions among teachers took place in an asynchronous way through the project website, where participants used forums, editable web shared documents (Google Drive) integrated in the website and specific forms for the completion of research activities.

The selection of participants for the activities conducted in the requirements phase was made through proposition to volunteers among teachers involved in Ledita project. The number of participants was 7, as suggested by GEM methodology, and we maintained this group along all the activities. The selected teachers come from Argentina, Brazil, Greece, Ireland and Italy and, as the others, are specialized in teaching Italian as second / foreign language. All the teachers had participated in previous LEDITA‟s research activities, usually carry out LD in their teaching practice and were previously introduced to the main software tools available for educational design and planning.

After discussion through an initial forum about some actual LD representations and tools, the teachers completed GEM activities in order to explicitly describe the concepts that characterize an ideal LD tool and to hierarchically classify these concepts. After reaching a consensus, results were critically discussed and teachers proceeded to the PAM activities. The aim of this second group of activities was: a) to elicit interested parties in the prospective software tool, with the Stakeholder Analysis artefact; b) anticipate possible problems and propose solutions, with the Evaluation Framing artefact and, finally, c) organize and discuss the results, highlighting eventually open issues, with the Semiotic Ladder artefact.

The next subsections describe the GEM as well as the PAM, their artefacts and the way they were used.

## 2.1 The Group Elicitation Method

The GEM (Boy, 1997) is a participatory practice we can locate in the initial stages of the software lifecycle, which aims at eliciting end-users" knowledge for the design of new user interfaces and complex human-machine systems. This participatory design method consists of the elicitation of important concepts from end-users" viewpoints and in deriving a consensus among the participants, using a brainstorming technique combined with a decision support system. A GEM session is usually composed by six phases: 1) Formulation of issue statements; 2) Generation of viewpoints; 3) Reformulation of viewpoints into more elaborate concepts; 4) Generation of relationships between these concepts; 5) Derivation of a consensus; 6) Critical analysis of the results.

The original phases of GEM were adapted to fit our research scenario in which the subjects had to participate at a distance, as follows:

Formulation of issue statements. For the formulation of issues statements, based on the list proposed by Nielsen et al. (1986), a structured interview was created and proposed to participants through a shared web document that teachers could simultaneously edit. The questions, translated into Italian, were as follows: What is the goal of the engineered system that we plan to design or evaluate?; How is the system or its equivalent being used (current practice, observed human errors); How would you use this system (users" requirements)?; What do you expect will happen if the corresponding design is implemented (e.g., productivity, aesthetics, quality of work product, quality of work life, and safety issues)?; How about doing the work this way (naive or provocative suggestions)?; What constraints do you foresee (pragmatic investigation of the work environment)?

Generation of viewpoints. This phase consisted of a “brainwriting”, a collaborative written brainstorming, aimed to highlight the points of view of the participants in relation to the questions posed in the structured interview. In this study the viewpoints were collected through their collaborative writing of a single document using Google Drive. The participation of the teachers in this activity lasted 3 days, with contributions and comments inserted directly into the shared document.

Reformulation of viewpoints into more elaborate concepts. For the elaboration of viewpoints into concepts, participants highlighted possible important concepts in the text and then analyzed and developed a list of concepts by means of combinations and divisions, always using collaborative writing through Google Drive.

Generation of relationships between these concepts. For the identification of relationships between concepts a form was created in the project website in which participants had to choose whether a concept was more important (+1), equally important (0) or less important (-1) compared to all other concepts mentioned. The objective of this artifact, called "triangular matrix", is to serve as decision system for the classification and organization of concepts obtained from previous stages.

Derivation of a consensus. For the derivation of consensus, a data analysis of each participant's matrix obtained during the phase 4 was carried out, by the creation of a global matrix of the scores assigned to the relationships between concepts. Starting from the global matrix it is possible to derive the consensus, which is expressed with 4 parameters: 1) The mean priority (MP) of a concept corresponds to the mean of the scores assigned to a concept with respect to the other concepts by all the participants. The value range of the mean priority is the interval [-100, +100]; 2) The interparticipant consistency (C) of a concept corresponds to the mean of the standard deviations of all global scores; 3) The mean priority deviation (D) or stability of a concept corresponds to the standard deviation of the mean priority with respect to the global scores of a concept; 4) The global consensus (GC) expresses a global score of the group consensus on the investigated issue.

Critical analysis of the results. Finally, results obtained from previous phases were presented to the participants, who have analyzed and commented on them using a forum in the website project.

## 2.2 The Problem Articulation Method

The PAM (Liu, 2000), developed in the later 1970s by Ronald Stamper within the MEASUR (Methods for Eliciting, Analyzing and Specifying Users‟ Requirements) research project, provides a set of techniques and tools that enable to understand and clarify problems. By using the method, undesirable omissions from analysis and specification can be reduced. Specifically, for the Ledita project, the same participants of previous GEM activities, always in remote and asynchronous activities, have used three artifacts: 1) Stakeholders Analysis, 2) Evaluation Framing, 3) Semiotic Ladder.

Stakeholder Analysis. This artifact allows investigating the involved parts that direct or indirectly influence or interest the information system under analysis. It is based on the technical, formal and informal levels of participation and organizes the stakeholders into five categories: Operation, Contribution, Source, Market and Community. To carry out this analysis, a document in Google Drive was prepared with the five stakeholders categories that participants filled in with their suggested stakeholders.

Evaluation Framing. The second activity consisted in completing the results obtained from the stakeholder analysis, by anticipating, for each stakeholder category, problems, questions and related issues and suggesting possible solutions (Baranauskas et al., 2005). For this activity, we prepared a Google Drive document with a table that, resuming the results of the Stakeholder analysis, added 2 columns to every stakeholder category: the first concerning problems/questions related to those stakeholders, and the second concerning ideas/solutions related to the raised issues.

Semiotic Ladder. To complete the PAM, participants filled in the Semiotic Ladder, an artifact useful to organize the different levels of requirement information. Besides the traditional semiotic division of syntax, semantics and pragmatics, the Semiotic Ladder of Stamper (1996) adds three new levels: "Physical World", "Empirics" and "Social World" (Table 1).

The activity of the participants in this study consisted in completing the various levels of the Semiotic Ladder starting from the stakeholders list suggested in previous analysis and indicating open questions and possible solutions for each level of the ladder. As for precedent activities, participants wrote their contributions directly in a web-shared document created with Google Drive.

**Table 1.** Original Semiotic Ladder (from Liu 2000)

<table>
<tr><td rowspan="3">Human information functions</td><td>SOCIAL WORLD: beliefs, expectations, culture, functions, commitments, contracts, law, ...</td></tr>
<tr><td>PRAGMATICS: intentions, conversations, negotiations, communications, ...</td></tr>
<tr><td>SEMANTICS: meanings, propositions, validity, truth, significations, denotations, ...</td></tr>
<tr><td rowspan="3">The IT platform</td><td>SYNTACTICS: formal structure, language, logic, data, records deduction, software, files, ...</td></tr>
<tr><td>EMPIRICS: pattern, variety, entropy, channel capacity, noise, redundancy, efficiency, codes, ...</td></tr>
<tr><td>PHYSICAL WORLD: signals, traces, physical distinctions, hardware, component density, speed, economics, ...</td></tr>
</table>

# 3 Results

Results from GEM and PAM activities were collected in text documents and spreadsheets and manually elaborated for analysis.

## 3.1 Findings on the Group Elicitation Method

For GEM activities, the teachers‟ participation was intense and every point of the issue statements was commented with the creation of articulated and connected viewpoints. Then, through several rounds of elaboration, 12 concepts were highlighted, interpreting and organizing the five-page document created in previous phases. The selected concepts were: 1) Support to Design, 2) Graphical Representation of Designs, 3) Consideration of Educational Needs, 4) Support to Reflection, 5) Economy (Time), 6) Ease of Use (Short Learning Curve), 7) Sharing of Designs, 8) Reuse of Designs, 9) Collaboration, 10) Author Identification, 11) Aesthetics (Look and Feel), 12) Software Compatibility.

In the generation of relationships between the concepts, we obtained a triangular matrix for every participant. By collecting the triangular matrix of all participants, we obtained the Global Score matrix (see Table 2). In this Table, the value of a single cell is related to the sum of all scores assigned by a participant in the triangular matrix to the relations of that concept with all the other concepts.

**Table 2.** Concepts Relationships Global Score Matrix

| Concepts | Single Participants' Score | | | | | | | Global Score (GS) |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| **Support to Reflection** | 7 | 3 | -3 | -1 | 6 | 4 | 7 | 23 |
| **Consideration of Educational Needs** | 7 | 2 | 1 | -5 | 6 | 4 | 6 | 21 |
| **Ease of Use (Short Learning Curve)** | 0 | 6 | -3 | 9 | 1 | 3 | 1 | 17 |
| **Economy (Time)** | 4 | 9 | -1 | 7 | -8 | -2 | 3 | 12 |
| **Reuse of Designs** | 2 | 4 | -2 | 6 | 3 | 4 | -5 | 12 |
| **Support to Design** | -2 | 2 | 3 | -5 | 2 | -1 | 8 | 7 |
| **Sharing of Designs** | 2 | -6 | 0 | 5 | 3 | 2 | -6 | 0 |
| **Software Compatibility** | 2 | 8 | -2 | -5 | 0 | 3 | -8 | -2 |
| **Graphical Representation of Designs** | -1 | -3 | 3 | -5 | -2 | -6 | 7 | -7 |
| **Collaboration** | -5 | -5 | -5 | 6 | 5 | 1 | -6 | -9 |
| **Author Identification** | -7 | -9 | 9 | -11 | -9 | -5 | -5 | -37 |
| **Aesthetics (Look and Feel)** | -9 | -11 | -1 | -1 | -8 | -7 | -2 | -39 |

Observing the Global Score obtained by each concept, we can notice that "Support to Reflection" and "Consideration of Educational Needs" were the most important concepts for participants, followed by "Ease of use (short learning curve)", "Economy (time)", "Reuse of designs" and "Support to Design", all with a positive score. "Sharing of designs" was understood as neutral, whereas "Software Compatibility", "Graphical representation of designs", "Collaboration", "Author identification", "Aesthetics (Look and feel of the software)" received a negative evaluation in relation to other concepts, meant to be less important.

In relation to the consensus analysis, results of Table 3 show that Mean Priority reflects the Global Score, with a sufficient uniformity of evaluation among the participants, except for the two most and, especially, the less important concepts, for which Mean Priority Deviation increases to exceed 5 points. These last values have therefore reduced the level of homogeneity, as we can see even from the relatively high Inter-participant Consistency value and from the Global Consensus that is slightly negative.

**Table 3.** Consensus Analysis

| Concepts | Mean Priority | Mean Priority Deviation |
|---|---|---|
| Support to reflection | 29,87 | 3,44 |
| Consideration of educational needs | 27,27 | 3,14 |
| Ease of use (short learning curve) | 22,08 | 2,54 |
| Economy (time) | 15,58 | 1,79 |
| Reuse of designs | 15,58 | 1,79 |
| Support to Design | 9,09 | 1,05 |
| Sharing of designs | 0 | 0 |
| Software Compatibility | -2,6 | 0,3 |
| Graphical representation of designs | -9,09 | 1,05 |
| Collaboration | -11,69 | 1,34 |
| Author identification | -48,05 | 5,53 |
| Aesthetics (Look and feel) | -50,65 | 5,82 |
| **Interparticipant Consistency: 4,34** | | |
| **Global Consensus: -2** | | |

## 3.2 Findings on the Problem Articulation Method

Results of the Stakeholder Analysis reported 19 stakeholders potentially interested in the system. Grouped in the categories of analysis, they are: Operation (Teachers, Students, Authors of didactical materials, Curriculum administrators), Contribution (Project participants, Software developers, Researchers), Source (Italian language teachers, Web system), Market (Universities involved in the research, LD software tools, Sponsors), Community (School directors, Faculty chairs, Media, Students‟ families, Teachers‟ families, Libraries, Scientific community). Figure 1 shows the results of stakeholder analysis in the Google Drive created by the participants.

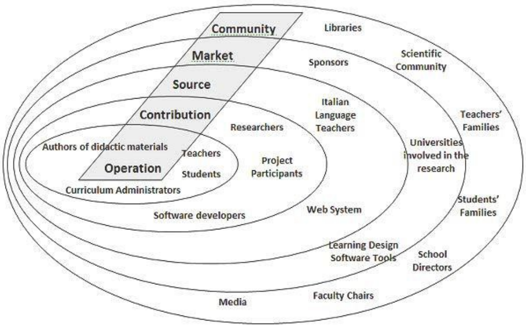



**Fig. 1.** Stakeholder Analysis

For the Evaluation Framing, participants filled in the predisposed form with 13 questions/problems and 11 ideas/solutions to solve these issues. Table 4 shows an excerpt of the Evaluation Framing structure, related to source stakeholders.

**Table 4.** Excerpt of Evaluation Framing

| Source | | |
|---|---|---|
| **Stakeholders:** | **Questions and Problems:** | **Ideas and Solutions:** |
| - Italian Language Teachers<br>- Web system | If the system is a web application, it needs an Internet connection to work. If a connection is not available, the system turns unusable. | Make available a system version that can be used without an Internet connection and provide the ability to upload the material developed offline when a connection is available. |

Finally, results of the Semiotic Ladder show a good participation for the three first levels (social world, pragmatics and semantics) and a more synthetic participation for the more technical levels (syntactics, empirics, physical world). Table 6 shows an excerpt of the Semiotic Ladder.

**Table 5.** Excerpt of Semiotic Ladder

| | **Elements:** | **Open Questions:** |
|---|---|---|
| **Social World** | Allow teachers' reflection on their teaching practice and facilitate a more efficient use of technologies in education | Could culture, values and emotions of teachers affect the use of software?<br>Are there laws that may create obstacles to the sharing of designs and resources?<br>How to ensure com- |

## 4 Discussion on the Main Findings

Results of the GEM activities showed essentially a strong interest of the participants in issues that are closely related to the practice of teaching. The main indication that comes from the elicitation and hierarchy of these concepts is the importance of LD as a moment of reflection and professional growth. This affirmation arose from the priority of “Support to Reflection” concept and from two explicit references in viewpoints for design practice as opportunity for professional growth. The supporting action is seen as a design flexibility that allows the “Reuse of designs” (appeared in concepts) and the revision and adaptation of designs to “Educational needs” (both in concepts and viewpoints).

This request for flexibility opens a new scenario in the horizon of LD tools to date, characterized by two main tendencies. The first is to support the design process via a user-friendly visual design environment, based on specific design principles and philosophies (Katsamani & Retalis, 2013); the second, is to help and guide teachers to take decisions during the design process (Masterman & Craft, 2013). However, teachers request a freer design process that is able to support and not constrain their ideas, choices and decisions. The flexibility is especially required by teachers' interest in ensuring the valorization of LD actors and resources. For the actors, the consideration of all possible subjects of an educational action is important, be they children, adults, and elderly or with special needs (four participants mentioned these during the generation of viewpoints). Concerning the resources, in facilitating the reuse and dissemination of educational materials previously created (three mentions in the viewpoints). The reuse of a LD is also motivated by the considerable amount of time required to design ("Economy" concept). In order to minimize this problem, a strong demand for usability and simplicity of the software emerged from teachers ("Ease of use" concept).

The importance attributed to the reflection on teaching practice and to the reuse of projects after a re-adaptation to the new context of use has contributed to the positioning of sharing of designs with other teachers in a secondary position ("Share of designs" concept collected MP = 0). This indication seems to go against the viewpoint of many LD experts, who argue that the sharing of designs between the community of practitioners is fundamental (Conole, 2008; Davinia Hernández-Leo et al., 2011). This is probably due to the fact that teachers have understood the sharing of designs as a not very useful activity if automatically done and not accompanied by reflection and the possibility to adapt the design to their needs (two teachers explicitly affirmed this in the viewpoints). This could also be due to the fact that there is no common language for describing online and face-to-face educational experiences (Dalziel, 2012).

Analyzing the viewpoints created by participants during GEM activities, we can highlight some interesting aspects. First, text emerged as a main representation modality for the design and graphical representations were limited to marginal roles. Furthermore, participants have always reported text editors as the main design tool that allows describing educational activities in detail. This indication contrasts with the current trend of LD software tools, for which, the representation is mainly graphical, using flowcharts, columns or concept-maps (Conole, 2013).

Little significance was given to the possibility of designing in a collaborative way (no explicit mention in the viewpoints and negative score for the mean priority of "collaboration" concept). This indication highlights the importance and uniqueness of the relationship between designer, educational context and teaching materials (reported in three different parts of the document). In this relation, the teacher/designer him/herself is seen as part of a system and not as the owner of an educational project (as shown by the negative score for the mean priority of the „author identification‟ concept). Another point is the possibility of taking into account the copyright rules for the use of specific educational resources. This element has been considered in GEM viewpoints, in the Evaluation Framing and in the Semiotic Ladder.

Concerning the more technical aspects, teachers have shown interest in a system that can adapt to multiple operating systems and devices. This request was made in order to make the system accessible by schools with poor technological structures. Finally, a lack of interest, although with a few exceptions, in the aesthetics of the software, reaffirming the need for simplicity and familiarity with the most common systems, especially text editors.

In relation to PAM activities, the stakeholder analysis has enriched the relationship between teachers, context and learning materials emerged from GEM, emphasizing the need to consider, in addition to teachers and educators, creators of educational materials, pedagogical coordinators and school managers. The presence of these stakeholders has led the discussion within the evaluation framing through aspects related to the management of the copyright for educational materials and the license to be applied to the software. The emerged intention, in line with current trends, was to move toward open materials and resources, allowing the interaction with the web for their retrieval, and to distribute the software with a free use license.

Another element of reflection was the difficult relationship of many teachers with technology, although they had considered themselves as knowers of technology in the first phase of the Ledita project. This difficulty appeared in relation to the use of tools other than those they are accustomed (office suite and graphic programs), and the frequent limitation of technological resources of the schools. In this regard, teachers stressed the importance of compatibility of the system with different devices, to provide simple and quickly visible instructions of use, and to generate a printing version of LDs, in order to facilitate the activities in the classroom, even in the absence of technological resources.

The Semiotic Ladder, finally, has encouraged a lively dialogue among teachers that has enriched previous discussions and has allowed analyzing elements of extreme importance for the development of the software. First, at the level of the social world, the reference to design as a tool for reflection on professional practice; efficiency in the use of technology in education has strengthened the demand for the development of an open system that makes the web a source of stimulation for the exchange and dialogue between cultures. This interpretation gives a new importance to the sharing of designs, which is not seen as an end, but as a motivation to the improvement and professional growth by means of the example and the re-elaboration of designs.

To this end, with the pragmatic level, the demand for flexibility of the software is found to be of fundamental importance for the success of the system. To allow adequate representation and fruitful sharing of designs, it is necessary a dynamic categorization of the elements that compose the designs. This could be achievable by allowing the customization by teachers, to suit their specific needs and better adapt to the educational context. In fact, one of the main limitations encountered by participants during the initial analysis of existing systems was the narrowness of some categorizations and the lack of possibility to add new elements.

Concerning flexibility and the good usability of the system, the semantic level brought the need to provide searching tools to explore all the possible design contents and combine the textual representation of the designs with a graphic summary that allows a global overview on the elements that compose the designs. Syntactic, empirical and the physical world levels have focused the attention on the development of a web-based system, that should allow safe access to users through a free registration service and the ability to be used on different types of devices, including desktop computers, laptops and tablets.

Summarizing, the main user requirements indications for the prospective LD tool, as resulted of this study, are listed as follows:

- Reflection and professional growth are the main aims for design practice and sharing. This indication requires the use of high level language and the selection of metaphors closer to teaching practice;
- Reuse of design is important for time economy and to stimulate the sharing of experiences, but only if designs can be modified and adapted to the new context of use;
- Flexibility is a key factor to adapt designs to every educational context. It is referred to the definition and orchestration of actors, resources and activities, using dynamic categorizations;
- Usability and simplicity are important for the diffusion among teachers;
- Text is the preferred modality of representation, whereas a graphical representation is useful for a global overview of the design;
- The system should be a web application to allow the use of different operational systems and devices;
- The use of free web resources allows avoiding copyright issues for didactical material.

## 5 Conclusion

LD is a key element to achieve positive educational results, but systems available today to support the LD process have not yet reached teachers and an adequate level of usability. This paper investigated the subject by conducting a study with teachers of Italian as second/foreign language, to understand the meaning they make to a prospective system intended to support their practices of LD.

This study involved the use of semio-participatory practices with a group of teachers at a distance, to understand the meaning they make to issues regarding LD practices and representations, aiming at eliciting user requirements for a prospective LD tool. The participatory requirement analysis activities carried out with the teachers have revealed aspects of the professional world of potential end-users and their needs and expectations. These participatory activities were well received by the participants and the remote asynchronous modality of participation has allowed us to complete the activities within a reasonable time and with a sufficient level of detail and involvement.

Analyzing the results, it was possible to synthesize a number of practical indications useful for developers interested in development informed by the practice of the main interested parties, who can rely on contextualized and well-argued information. Future works in this investigation involve the formalization of a conceptual framework able to support reflection and professional growth within the practice of educational design and the development of a system capable to respond to the user requirements emerged from this study.

## References

1. Agostinho, S.: The use of a visual learning design representation to document and communicate teaching ideas. In: Proceedings of the 23rd Annual ASCILTE Conference: Who's Learning? Whose Technology?, Sidney, pp. 3–7 (2006)
2. Arpetti, A., Baranauskas, M.C.C., Leo, T.: Making Design Easy: a Usability Evaluation of Latest Generation Learning Design Tools. In: Herrington, J., Couros, A., Irvine, V. (eds.) World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA), pp. 960–965. AACE, Victoria (2013a), http://www.editlib.org/p/112076 (retrieved)
3. Arpetti, A., Baranauskas, M.C.C., Leo, T.: Learning Design and Teaching Practice: Outlining an Iterative Cycle for Professional Teachers. In: 2013 IEEE 13th International Conference on Advanced Learning Technologies, pp. 280–284. IEEE (2013b), doi:10.1109/ICALT.2013.87
4. Baranauskas, M.C.C.: Socially aware computing. In: ICECE 2009 VI International Conference on Engineering and Computer Education, pp. 1–5 (2009)
5. Baranauskas, M.C.C., Schimiguel, J., Simoni, C., Bauzer Medeiros, C.: Guiding the Process of Requirements Elicitation with a Semiotic-based Approach. In: 11th International Conference on Human-Computer Interaction, pp. 100–111. Lawrence Erlbaum, Las Vegas (2005)
6. Beetham, H.: An approach to learning activity design. In: Beetham, H., Sharpe, R. (eds.) Rethinking pedagogy for a digital age: Designing and delivering e-learning. Routledge, London (2007)
7. Beetham, H., Sharpe, R. (eds.): Rethinking Pedagogy for a Digital Age: Designing for 21st Century Learning. Routledge, New York (2013)
8. Boy, G.A.: The group elicitation method for participatory design and usability testing. Interactions, 27–33 (1997)
9. Chu, S.K., Kennedy, D.M.: Using online collaborative tools for groups to coconstruct knowledge. Online Information Review 35(4), 581–597 (2011)
10. Compendium Institute (2013), http://compendium.open.ac.uk/institute/ (retrieved January 05, 2014)
11. Conole, G.: Tools and resources to guide practice. In: Rethinking Pedagogy for a Digital Age: Designing for 21st Century Learning, pp. 1–25 (2008)
12. Conole, G.: An overview of design representations. In: Proceedings of the 7th International Conference of Networked Learning (NLC 2010), Aalborg, Denmark, pp. 482–489 (2010)
13. Conole, G.: Designing for Learning in an Open World. Springer, Berlin (2013)
14. Conole, G., Weller, M.: Using learning design as a framework for supporting the design and reuse of OER. Journal of Interactive Media in Education 2008(1), 1–13 (2008)
15. Dalziel, J.R.: Lessons from LAMS for IMS Learning Design. In: Sixth IEEE International Conference on Advanced Learning Technologies (ICALT 2006), vol. (3), pp. 1101–1102 (2006)
16. Dalziel, J.R.: The Larnaca Declaration on Learning Design (2012), http://www.larnacadeclaration.org
17. Derntl, M., Neumann, S., Oberhuemer, P.: Propelling Standards-based Sharing and Reuse in Instructional Modeling Communities: The Open Graphical Learning Modeler (OpenGLM). In: 2011 IEEE 11th International Conference on Advanced Learning Technologies, pp. 431–435. IEEE (2011)
18. Dobozy, E.: Typologies of Learning Design and the introduction of a "LD-Type 2" case example. eLearning Papers 27(27), 1–11 (2011)

19. Earp, J., Pozzi, F.: Fostering reflection in ICT-based pedagogical planning. In: Proceedings of the First International LAMS Conference, pp. 35–44 (2006)
20. Ferrell, G.: Transforming curriculum design: transforming institutions, pp. 1–4. Briefing (February 2011)
21. Goodyear, P.: Educational design and networked learning: Patterns, pattern languages and design practice. Australasian Journal of Educational Technology 21(1), 1–12 (2005)
22. Goodyear, P., Ellis, R.: Students interpretations of learning tasks: Implications for educational design. In: Proceedings of the ASCILITE 2007 Conference, pp. 339–346 (2007)
23. Griffiths, D., Beauvoir, P., Sharples, P.: Advances in Editors for IMS LD in the TENCompetence Project. In: 2008 Eighth IEEE International Conference on Advanced Learning Technologies, pp. 1045–1047. IEEE (2008)
24. Gutierrez, S., Valigiani, G., Jamont, Y., Collet, P., Kloos, C.D.: A Swarm Approach for Automatic Auditing of Pedagogical Planning. In: Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007), pp. 136–138. IEEE (2007)
25. 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)
26. Katsamani, M., Retalis, S.: Orchestrating learning activities using the CADMOS learning design tool. Research in Learning Technologies Supplement 21, 1–12 (2013)
27. Katsamani, M., Retalis, S.: Making learning designs in layers: The CADMOS approach. In: Proceedings of the IADIS Multi Conference on Computer Science and Information Systems, pp. 305–312 (2011)
28. Koper, R., Olivier, B., Anderson, T.: IMS learning design information model. IMS Global Learning Consortium (2003)
29. Koper, R.: Current Research in Learning Design. Educational Technology & Society 9(1), 13–22 (2006)
30. Liu, K.: Semiotics in information systems engineering. Cambridge University Press, Cambridge (2000)
31. Lockyer, L., Bennett, S., Agostinho, S., Harper, B.: Handbook of Research on Learning Design and Learning Objects: Issues, Applications, and Technologies, vol. 2. IGI Global, Hershey (2009)
32. Masterman, E., Craft, B.: Designing and evaluating representations to model pedagogy. Research in Learning Technologies Supplement 21, 1–14 (2013)
33. McKenney, S., Reeves, T.C.: Conducting Educational Design Research, p. 234. Routledge, New York (2012)
34. Nielsen, J., Mack, R.L., Bergendorff, K.H., Grischkowsky, N.L.: Integrated software usage in the professional work environment: evidence from questionnaires and interviews. ACM SIGCHI Bulletin 17, 162–167 (1986)
35. Oliver, R., Littlejohn, A.: Discovering and describing accessible and reusable practitioner-focused learning. In: Proceedings of Theme of the JISC Online Conference: Innovating e-Learning, pp. 30–33 (2006)
36. Persico, D., Pozzi, F., Anastopoulou, S., Conole, G., Craft, B., Dimitriadis, H.D., Kali, Y., Mor, Y., Pérez-Sanagustín, M., Walmsley, H.: Learning design Rashomon I through different approaches. Research in Learning Technologies Supplement 21 (2013)
37. RELOAD Reusable eLEarning Object Authoring & Delivery (n.d.), http://www.reload.ac.uk/ (retrieved January 05, 2014)
38. Stamper, R.: Signs, Information, Norms and Systems. In: Holmqvist, B., et al. (eds.) Signs at Work, pp. 349–397. Walter de Gruyter, Berlin (1996)
39. Wijnen, W.: Towards Design-Based Learning. Eindhoven: Eindhoven University of Technology. University of Eindhoven, NL: OGO brochure, No2. Educational Service Center (2000)

# Analysing Microblogs of Middle and High School Students

Javiera Born[1], Eliana Scheihing[2,⋆], Julio Guerra[3], and Luis Cárcamo[4]

[1] Universidad de Chile, Santiago, Chile
jborn@dcc.uchile.cl
[2] Universidad Austral de Chile, Valdivia, Chile
escheihi@inf.uach.cl
[3] Universidad Austral de Chile & University of Pittsburgh, Pittsburgh, USA
jguerra@inf.uach.cl
[4] Universidad Austral de Chile, Valdivia, Chile
luis.carcamo@gmail.com

**Abstract.** The Kelluwen project implements middle and high-school educational activities where the use of Web 2.0 tools is incorporated to improve collaboration construction, sharing and publishing of the learning outcomes. The Worklog tool, a microblogging space within the Kelluwen platform has an active role in the educational activities. Using probabilistic topic models, correlation analysis and principal component analysis (PCA), we analyzed micropost of 85 class groups participating in the Kelluwen project and found interesting relations of the types of messages posted and other factors such as the teacher participation in the microblog, the rural or urban nature of the schools and other aspects of the educational experience.

## 1 Introduction

A number of public educational policies and efforts have been carried out in Chile to incorporate technologies in the public educational system [9,2]. In the decade of the 90', the educational policies implemented the project Enlaces (www.enlaces.cl) which successfully deploy computer laboratories to nearly 100% of all primary and secondary schools in the country. Nowadays, this infrastructure is maintained and used in different levels, and in many cases, with a considerable amount of difficulties. In many cases, the school administrators set low importance to computer maintenance, networking and Internet connection, and reduce costs on the maintenance of this infrastructure. In other cases, due to disciplinary and educational concerns, some schools block the access to on-line systems like Facebook or Youtube, because they are considered as distracting resources. In summary, administrative difficulties, and cultural and technological gaps on teachers and administrative staff join together to make the incorporation of educational innovation using Web 2.0 in the school a very difficult endeavor

⋆ Financial support from Fondef- CONICYT under D08i-1074 project and DID-UACh.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 15–28, 2014.
© Springer International Publishing Switzerland 2014

[4]. In the other hand, the use of the Internet has grown very fast in Chile. Consumption studies have identify that 91% of the young population and 56% of the adults and elder adults are connected to the Internet. Thus, the use of social networking on-line tools and social media has reached the 89% of the population ref:iab2013.

The project Kelluwen [14] is an attempt to incorporate the effective use of social web tools (the Web 2.0) in formal curriculum of public education in southern Chile. The project target the socio-communicative skills in high school students and put forward activities combining traditional classroom activities and the use of social web tools as didactic resources. To better support such activities, the Kelluwen Web platform includes a microblogging tool called the Worklog (Bitácora in spanish) [12]. The Worklog tool was originally added as a complementary and informal channel for students and teachers to share information within and between different classrooms participating in the project while performing similar activities. During activities, students and teachers used the worklog to give opinion on the students' team works, share information, reflect about the ongoing activities, or simply say "hello". We claimed that the worklog is a valuable tool since it's logs of messages can be further reviewed to partially reconstruct or have a better idea of what "was going on" during the activities. Since the worklog was explicitly intended as an informal channel, one of the challenges of analysing the microblogging activity is to deal with the heavy use of slang and miswritten words.

In this work we use statistical based topic models to explore the microblogging messages of 85 classroom groups that used the worklog tool during 2011 and 2012, relating such activity with other parameters such as the participation of teachers, the level of completeness of the activities, and the urban/rural nature of the schools. We are interested to see to which extent the participation of the teachers in the worklog influences students' posts and to which extent the activity on the worklog reflect the success of the learning experiences.

The paper is organized as follows. The next section briefly reviews previous work done in two areas: text analysis on microblogging data, and the use of microblogs in educational settings. Section 3 describes the educational setting carried out in Kelluwen interventions, the microblogging platform -the Worklog- that support the activities, and the data set collected. Section 4 presents the pre-processing of the data set, the variables included in the analysis and the results of the topic model. Section 5 describes the results of the analysis of the use of the microblogging space and the factors that quantitatively and qualitatively describe the learning experiences. Section 6 presents the conclusions and discussion.

## 2 Related Work

### 2.1 Microblogging in Education

Regarding the use of Microblogging In Education (MIE), Gao et al [7] made a survey analysis of 21 research works published during the period of 2008 and

2011. They reported a wide variety of educational settings using microblogging, mainly as a informal communication platform supporting the learning activities. According to the authors, the analyzed works consistently concluded that microblogging effectively increases communication (both between students and between students and teachers), increase social presence, and help in building a learning community. However, published work is usually descriptive and lacks of deep analysis on the impact of using microbogging in the class. According to Ebner et al [6], the use of microblogging has two main advantages: it makes possible to post immediate feedback (especially in peer review and collaborative learning activities) and it allows to track the learning process. In their case study, students of a master degree program of Business in Austria used a microblog based on the open source software Indenti.ca. Other works report the impacts of using microblogging in reflective thinking [16,5]: as students are limited to 140 characters to express ideas, they are forced to express themselves in a concise and clear manner. Holotescu and Grosseck [10] showed how an entire course can be managed and delivered using a microblogging platform. They use Cirip.ro, a platform that enhance microblogging with group management, embbeded images and video, among other features.

### 2.2 Analysis of Microblogging Data

There is a considerable amount of research analyzing twitter (and other microblogging systems) activity in different context. In general, microblogging analysis is performed with a wide range of methods involving machine learning and natural language processing techniques. Abel *et al* [1] performed hashtags, entity, and topic extraction techniques using OpenCalais (www.opencalais.com) to build user profiles from twitter data. Paul and Drezde [13] presents the Ailment Topic Aspect Model (ATAM) where words describing ailments and diseases are taken from articles and used as prior knowledge to enhance a probabilistic topic modeling process done over twitter data covering public health trends. Yang *et al* [17] put forward a framework to perform dynamic pattern analysis on Twitter feeds. Their algorithm, SPUR (Summarization via Pattern Utility Ranking) analyzes topic patterns as the tweet feed goes on. Topics are extracted as a result of matrix factorization techniques, and divergence metrics are computed between topics in different times to build an evolution trace distinguishing topics that change, new topics appearance, topic disappearance, topic merging, etc.

## 3 The Kelluwen Activities and Platform

The Kelluwen project developed a series of *Didactic Designs* to be used in middle and high school levels in 2 classes: Language and History. A *Didactic Design* (DD) covers a unit of content as specified in the chilean public school curriculum, and contains a series of activities which blend traditional classroom learning strategies with the use of social web tools like Wordpress, Youtube, Panoramio, etc. In this way, Kelluwen aims to bring the sharing and communication capabilities of the social web, to formal learning activities in schools. A common DD sequences around 9 activities, and each activity is intended to last 1 class period (1 and a half hours). In the DD, the students work in teams with the goal of creating an outcome that is finally published in a social Web tool and reviewed by peers. To make this happen, activities are organized into three stages: a *motivation stage*, where students are introduced to concepts and topics, and they must study the topic and begin to incorporate the knowledge; a *creation stage*, where using the learned concepts they must generate new content, such as a blog or a video, exercising their communicative skills; and an *evaluation stage*, where the generated content is now passed to other students, which may be from the same classroom or from a different classroom working in the same DD, who must evaluate and give feedback on the work done. They also have the opportunity to do an auto-evaluation and receive comments from the guiding teacher.

During the activities, students and teachers have access to the Kelluwen Platform, a Web system designed to support the learning activities by providing access to material and tools to facilitate the communication and monitoring of the work. Although the main interaction between students and teachers happens in person, the Kelluwen Platform and its tools are used to a considerable extent. As mentioned before, the platform include a tool called the Worklog that provides each class with a microblogging space. Within the 3 years of the Kelluwen project interventions, the use of the Worklog tool varied as the project developed and other tools for communication and work management were added to the platform. Until the first semester of 2011, the microblog was mainly used to share links to the teams work outcomes and was the only communication tool between different classrooms. Up to that moment, the use of the Worklog was generally mandatory in most of the classrooms. After a new tool for work publishing and peer review [15] was deployed to the Kelluwen platform in the second semester of 2011, the worklog tool started to be used as a complementary non-mandatory tool. Even when optional, students tend to largely use it for greetings, sharing information, commenting and giving opinions. Additionally, and since it is an informal communication space, the Worklog is full of slang, emoticons (expressions of emotions through ASCII characters) and miswritten words, which make challenging to perform any text analysis of the posts.

### 3.1 The Dataset

We collected microblogging posts of 85 class groups working on 24 different Didactic Designs (DD) during the second semester of 2011 and the first semester of 2012 as during these periods the Worklog use was optional (we believe that the analysis of posts when the tool was of non-mandatory use is more relevant). Each DD is designed for one of two areas: Language or History. There were 46 class groups working on DDs in the area of Language, and 39 in the area of History. Interventions covered middle and high school: there were 48 class groups from levels 6-8, and 37 class groups from levels 9-12 (in the chilean educational system, the grades 9-12 correspond to high school). Participating schools are from different cities and towns covering great part of the southern area in Chile. Some cities are relatively large urban zones and some cities are middle or small size cities dominated by rural characteristics. We labeled all school as urban or rural depending on the size of the city. There are 64 class groups from schools located in urban areas, and 21 class groups from rural areas. About the messages in the worklog, there are a total of 9010 posts, where 8432 were posted by 1849 students (there were 2540 registered students organized in 576 teams) and 510 were posted by teachers. Histograms of the students and teachers posts by class group are showed in Figure 1, top and bottom charts, respectively. It is surprising that many students posted since, as a blended learning activity, the access to a computer and Internet connection during classes was in many cases restricted to one computer per group.

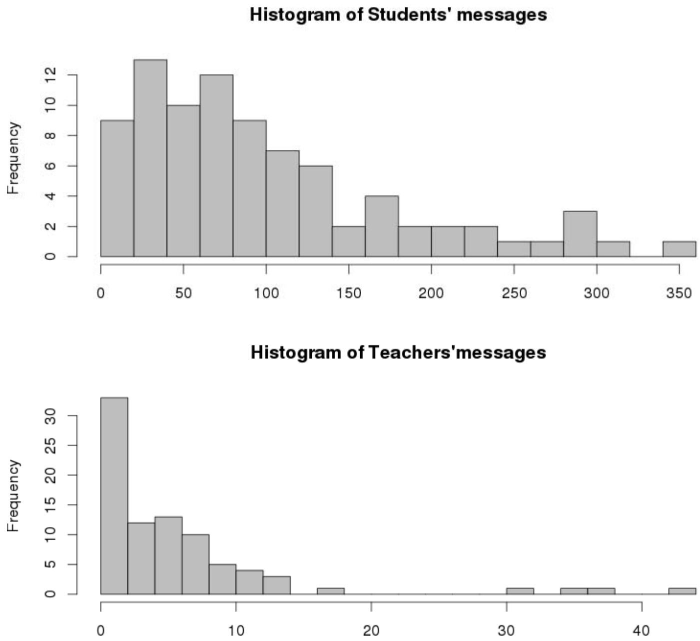



**Fig. 1.** Histograms for the number of messages (posts) from students (top) and teachers (bottom). The x-axis represents the number of posts. Students' posts distribution has a median of 74.0, mean of 99.2, and 50% of data between 38 and 127, and a maximum of 343 messages in one of the class groups. Teachers' distribution has a median of 4, mean of 6, and a maximum of 43 messages in one of the class groups.

# 4 Analysis of the Microblogging Posts

## 4.1 Preparing the Data

A first general observation of the posts shows that most repeated words were a) emoticons, that is, expressions of emotions through ASCII characters,

b) links to videos, images and articles on the web, and c) connectors of sentences, both well written and mis-written. For a more significant result, these 3 aspects were partially removed from the messages. Both a) and c) presents a set of words which should be removed, along with numbers and punctuation. Words with 1 character or more than 20 were removed as well, as they don't provide information. For b), all links containing *http://* were substituted by only *http*, with the exception of Youtube videos which were substituted by the *youtube* string (some didactic designs ask for publishing videos in Youtube, although a special tool is available in the platform to put the links to the work done). We use R and the package *tm*[1] to perform these data filtering.

### 4.2 Topic Extraction

We are interested on extracting semantic features from the text of the messages and further contrast these semantic features with other parameters of the activity of the class groups like the level of participation of the teachers. After considering various options for the analysis of the microblogging messages, we leaned towards the Probabilistic Topic Models strategy using Latent Dirichlet allocation (LDA) [3]. LDA is widely used in text analysis to uncover latent topics within collections of text documents. The general assumption in LDA is that each document (in our case a document is a microblogging post) is generated from a latent set of topics, and each topic is a probability distribution among the entire vocabulary of the collection. LDA analyses co-occurrence of words in the documents to generate the set of topics (the number of topics should be defined in advance) by following two optimization goals: minimize the number of probable topics related to each document (describe each document as concisely as possible), and minimize the number of highly probable words within each topic (make topics as specific as possible). Observe that these two optimization goals are in conflict due to the multi-topic nature of the documents. Working with LDA has two main advantages. The first advantage is that the model can deal with multi-topic documents. In our case, a preliminary observation showed that roughly, microblogging posts are a combination of greeting (say hello in different ways), opinions (mostly when a team opines about the work of others), and specific information (providing a link or talking about the content related to the DD). The second advantage is that by grouping words that occurs together, LDA is able to deal with synonyms (different words with the same meaning) and polysemy (a word having different meaning depending on the context), which is of great use in our case where we have a considerable amount of mis-written words.

As mentioned before, a preliminary observation roughly showed 3 topics: opinion, greeting, information. Considering this, we run the topic modeling for 3, 4 and 5 topics using *R* and the package *lda*[2] with 50000 iterations. We repeat each experiment 3 times and compute log-likelihood. Results of maximum

[1] `http://cran.r-project.org/web/packages/tm/vignettes/tm.pdf`
[2] `http://cran.r-project.org/web/packages/lda/lda.pdf`

log-likelihood run for the experiment with 3 topics are shown in the table 1 where the top 10 probable words within each topic are shown.

A review of the probabilities within each topic reveals that the top 10 words cover 12.12%, 14.42% and 10.81% percent of the topic, respectively. Note in the table that topic 1 and topic 2 are clearly about opinion and greeting, respectively. Topic 3 is informational and contains words related to the content of the DDs. When repeating the process for 4 and 5 topics, the first two topics (opinion and greeting) remain, and the rest of the topics split the words accordingly to the different content in different DDs. Because we aim to analyze the topics regardless of specific content of the didactic designs, we decided to consider only three topics in further analysis, where the third topic includes information messages from the different themes. We further refer to the topics by their general description: "opinion", "information" and "greeting". As we will see in section 5, we used those 3 topics and the topic-micropost probabilities to aggregate those probabilities to the level of the class group. In this form, we computed an *opinion* score, and *information* score and a *greeting* score for each of the 85 class groups.

**Table 1.** Original words and an approximate translation, where the three topics can be observed

| **Topic 1** | **Topic 2** | **Topic 3** |
|---|---|---|
| trabajo *(work)* | hola *(hello)* | mundo *(world)* |
| bueno *(good)* | ola *(hello)* | literario *(literary)* |
| gusto *(like)* | wena *(cool)* | texto *(text)* |
| actividad *(activity)* | cabros *(pals)* | textos *(text)* |
| trabajos *(works)* | oli *(hi)* | literarios *(literary)* |
| bien *(nice)* | gusta *(like)* | onirico *(oniric)* |
| parecio *(thought)* | saludos *(greetings)* | http |
| kelluwen | chicos *(boys)* | hechos *(facts)* |
| entretenido *(fun)* | gay | cotidiano *(daily)* |
| buena *(cool)* | holaa *(hello)* | vida *(life)* |

### 4.3 Indicators

The analysis performed in the next section contrast the topics extracted from the messages within a class group posts with various factors considered as indicators of the success of the activities. These factors are described below.

- **Participation of teachers:** we consider the fact that teachers participated in the microblog can influence the type of comments posted by students, especially pushing students towards opinion and informational messages. This indicator is computed as the percentage of teacher messages over all messages posted in each class group and is named *teacher participation.*
- **Progress of the experience:** some class groups never finished the activities of the DD. This is a clear result of whether the teacher decided to use the platform or not, and if they did not, some influence related to the stage where the activity was abandoned might be visible in the microblog. This indicator is computed as the percentage of completed activities in each class group and is called *progress.*

- **Regularity of the activities:** related to the previous one, we would like to know if the regular use of the platform influence the messages. This is measured through obtaining the differences in days between activities, and determining the mean and variance for each class group, so that a small mean and variance implies regularity. The two factors are called *mean* and *variance.*
- **Number of Students:** we include in the analysis the number of students that belong to each class group, that is not necessarily the same number of students who participate in posting in the microblog. This factor is named *students.*
- **Location:** the schools are characterized as urban or rural depending on the size and importance of the city in which the school is located. Regional capitals are characterized by urban dynamic, and in contrast non-capital cities and towns have a strong dominance of rural aspects. This factor is named *location.*
- **Didactic Design:** although not a measured variable, we consider the Didactic Design as a factor in further analysis because it could explain the nature of the use of the microblogging tool. This factor is called *design.*

## 5 How the Micropost Topics Relate to the Class Indicators?

Using the 3 topic model obtained in the previous section, we aggregate the topic probabilities distributions of the posts belonging to the each class group. We compute class-level topic distributions by this aggregation approach since it has been pointed out that in the use of LDA to generate aggregated topic distributions, a good strategy is to input the documents as an unstructured collection and then perform a hierarchical aggregation of the topic distributions to the upper level [8]. In the following analysis we first compute correlations among the topic scores (*greeting*, *opinion* and *information* scores) and the other 7 indicators previously described. Then we perform Principal Component Analysis (PCA) to show the relation of the factors in a 2 dimensional visualization.

**Table 2.** Correlation matrix for 10 components

| | I | O | G | T | P | S | M | V | L | D |
|---|---|---|---|---|---|---|---|---|---|---|
| Information | | 0.42 | **-0.69** | 0.36 | 0.17 | -0.10 | 0.07 | 0.03 | -0.13 | 0.14 |
| Opinion | 0.42 | | -0.36 | -0.04 | 0.07 | 0.04 | -0.09 | -0.02 | 0.14 | -0.26 |
| Greeting | **-0.69** | -0.36 | | -0.34 | -0.23 | 0.07 | 0.00 | -0.02 | 0.02 | 0.07 |
| Teacher Part | 0.36 | -0.04 | -0.34 | | -0.23 | -0.10 | 0.48 | 0.17 | 0.19 | 0.26 |
| Progress | 0.17 | 0.07 | -0.23 | -0.23 | | 0.21 | -0.37 | 0.04 | 0.01 | -0.15 |
| Students | -0.10 | 0.04 | 0.07 | -0.10 | 0.21 | | -0.14 | -0.11 | 0.02 | -0.10 |
| Mean | 0.07 | -0.09 | 0.00 | 0.48 | -0.37 | -0.14 | | **0.66** | 0.05 | 0.29 |
| Variance | 0.03 | -0.02 | -0.02 | 0.17 | 0.04 | -0.11 | **0.66** | | 0.17 | 0.14 |
| Location | -0.13 | 0.14 | 0.02 | 0.19 | 0.01 | 0.02 | 0.05 | 0.17 | | -0.03 |
| Design | 0.14 | -0.26 | 0.07 | 0.26 | -0.15 | -0.10 | 0.29 | 0.14 | -0.03 | |

### 5.1 Correlation Matrix

Our first approach to the results is to observe the correlation matrix, which can be observed in Figure 2. It can be noted that correlation score ($x$) appear as significant in only two relations ($|x| \geq 0.5$): Between *Greeting* and *Information* scores, and between *Mean* and *Variance* of the regularity of the activities. The second relation is no surprise: there are few values on which these variables are calculated, and therefore, few outliers produce high variance. A *big* mean implies a big 'jump' in the times in the activities, and therefore, a very different value than the other small 'jumps', generating a high variance.

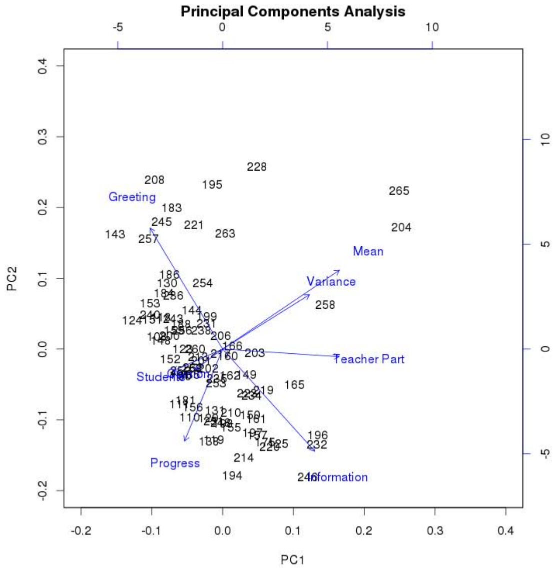



**Fig. 2.** Analysis over the 3 topic model probabilities and the indicators described. The main plane represents 40.08% of the data variance.

### 5.2 Principal Components Analysis

For a graphical analysis, we generate a Principal Component Analysis (PCA) over the data. The Figure 2, shows a two dimensional representation where correlations can be clearly observed as the arrow point to the same (positive correlations) or opposite (negative correlations) directions. The first correlation observed is the correlation between the *Mean* and *Variance*, previously described, and between *Greeting* and *Information* scores. A more interesting correlation is also observed in this chart: *Teacher Participation* seems to influence the probability of the messages of being labeled as *Information*. *Opinion* score, on the other hand, is inversely correlated to the *Mean* and the *Variance*, implying that a regular development of the experience influences the tendency to post *Opinion* messages. It can be noted that, in the top words of *Information* topic, the word

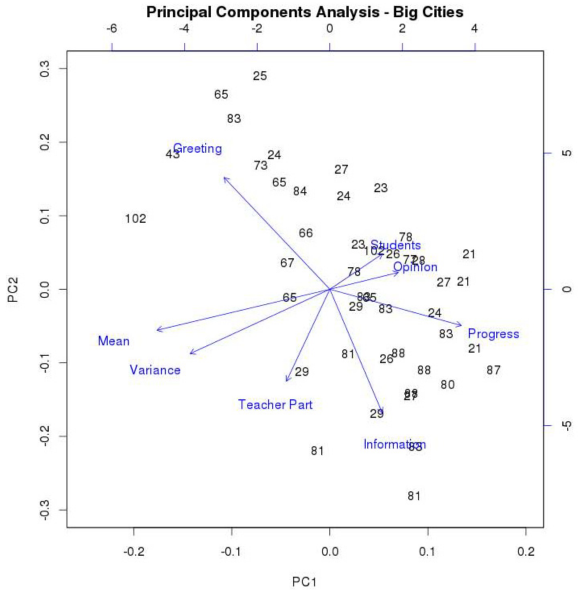



**Fig. 3.** PCA considering only class groups in urban schools. The main plane represents 41.00% of the data variance.

*http* appears, suggesting the relevance of the "creativity stage" where they may post more links. We are also interested in visualize the difference between the rural and urban school class groups, and therefore we break the PCA analysis into 2 groups separating the data by the *Location* factor. We present the results of our analyses in Figures 3 and 4. We focus on correlations between one of the three topics —Opinion, Information or Greeting— and the other indicators.

In both cases we observe positive correlations between *Mean* and *Variance*– correlation of 0.69 and 0.75 (urban and rural, respectively)— and negative correlations between *Greetings* and *Opinion* scores –correlation of -0.37 and -0.36 (urban and rural, respectively)— and between *Greeting* and *Information* –correlation of -0.70 and -0.69 (urban and rural, respectively)—. In contrast, different correlations are founded in the following cases:

1. **Urban school:** positive correlations between *Progress* and *Information*– correlation of 0.28— and a weak correlation of 0.16 between *Teacher Participation* and *Information* score, while negative correlations can be observed between *Greetings* and *Progress* —correlation of -0.31—.
2. **Rural schools:** positive correlations between *Teacher Participation* and *Information* score —correlated by 0.57—, between *Teacher Participation* and *Mean* of regularity —with a correlation of 0.55—, and inverse correlations between *Teacher Participation* and *Progress* —with -0.34 correlation.

These results suggest that in both cases (urban and rural) greeting messages take the place of other more advance use of the microblogging space (opinion

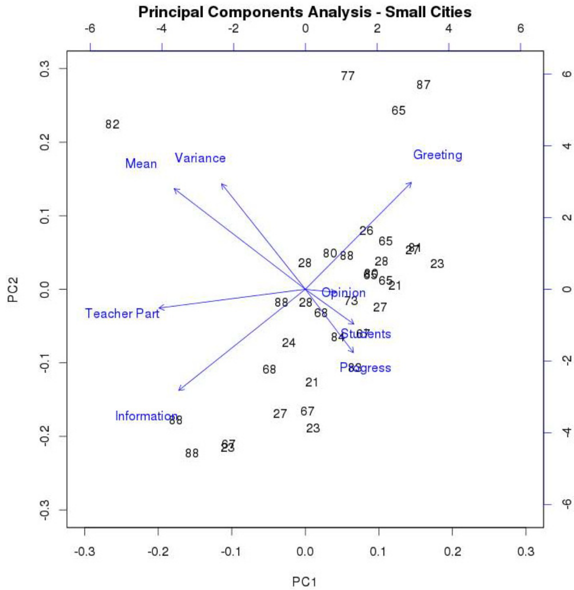



**Fig. 4.** PCA considering only class groups in rural schools. The main plane represents 40.22% of the data variance.

and information). While *Teacher Participation* relates to *Information* messages in rural areas, in urban areas this relation is weaker. In that case we observe a positive relation between *Information* score and Progress.

## 6 Conclusion

In this paper we conducted an analysis of microblog posts of middle and high school students in classroom participating in blended (face-to-face and virtual) activities. The Kelluwen project provides didactic designs in different subjects (language, history) which guide teachers and students in didactic activities incorporating social web tools. The Kelluwen platform includes a microblogging tool that students use freely to know each other, share resources or give feedback.

The results obtained with the LDA topic model showed that the messages posted by the students in the micro-blogging tool can be effectively grouped in three main topics: *opinion*, *greetings* and *information*. The information topic is related to the specific content of the different didactic design and this topic is splitted when the process run for higher number of topics: we tested from three to five topics and in all cases we found two distinctive topics corresponding to opinion and greeting messages, while the other topics corresponded to the information of the matters addressed by didactic designs. Because we performed a global analysis considering didactic designs from different subject areas, we decided to consider only three topics in our analysis, where the third topic includes information messages from the different content themes.

With the obtained topics we computed empirical probabilities that a message belongs to one of the three topics and then we related these probabilities with indicators describing each class group. The Principal Component Analysis showed that the degree of *Teacher Participation* positively influences the probability of *Information* messages. *Opinion* messages, on the other hand, are inversely correlated to the *Mean* and the *Variance* of the times between activities, implying that a regular development of the experience influences the tendency to post *Opinion* messages. When we analysed the behaviour of urban class groups, we observed that *Opinion* score is positively correlated to *Progress*, entailing that well executed experiences imply that the students produce more opinion messages, which is probably associated to the reflection about their learning processes.

These results confirm the potentiality of the worklog in various senses. On one hand, we have seen how the worklog is an effective tool for teacher mediation in the learning process: students post more about information or knowledge issues when the teacher is more active in the worklog. On the other hand, if we interpret the *opinion* messages as the results of motivation and reflective processes on the learning activities, we can explain the correlation between *progress* and *opinion* scores, considering that the progress in the learning activities is modulated by the motivation of students.

Finally, results highlight the value of communicative aspects in b-learning experiences: although the teachers have the opportunity to give offline instructions, their participation in the worklog tool establishes a closer communicative relationship with the students (dialogic and horizontal). This can promote motivational aspects and improve the quality of pedagogic scaffolding.

## References


1. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Analyzing user modeling on twitter for personalized news recommendations. In: Konstan, J.A., Conejo, R., Marzo, J.L., Oliver, N. (eds.) UMAP 2011. LNCS, vol. 6787, pp. 1–12. Springer, Heidelberg (2011)
2. Arancibia Herrera, M.: La Red Enlaces en Chile. Incorporación de la Informática a la Tareas Pedagógicas: una Historia que Alumbra un Promisorio Futuro. Perspectiva Educacional 36, 133–151 (2000)
3. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. The Journal of Machine Learning Research 3, 993–1022 (2003)
4. Crook, C.R.: The "digital native" in context: tensions associated with importing Web 2.0 practices into the school setting. Oxford Review of Education 38(1), 63–80 (2012), `http://dx.doi.org/10.1080/03054985.2011.577946`
5. Ebner, M., Maurer, H.: Can weblogs and microblogs change traditional scientic writing? Future Internet 1(1), 47–58 (2009), doi:10.3390/1010047.
6. Ebner, M., Lienhardt, C., Rohs, M., Meyer, I.: Microblogs in Higher Education - A chance to facilitate informal and process-oriented learning? Computers & Education 55(1), 92–100 (2010)
7. Gao, F., Luo, T., Zhang, K.: Tweeting for learning: A critical analysis of research on microblogging in education published in 2008-2011. British Journal of Educational Technology 43(5), 783–801 (2012)

8. Guerra, J., Sosnovsky, S., Brusilovsky, P.: When One Textbook Is Not Enough: Linking Multiple Textbooks Using Probabilistic Topic Models. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 125–138. Springer, Heidelberg (2013)
9. Hepp Kushel, P.E.: ENLACES: El Programa de Informática Educativa de la Reforma Educacional Chilena. In: Cox, C. (ed.) Políticas Educacionales en el Cambio de Siglo. La Reforma del Sistema Escolar de Chile, pp. 419–451. Editorial Universitaria, Santiago (2003)
10. Holotescu, C., Grosseck, G.: Using microblogging in education. Case Study: Cirip.ro. In: 6th International Conference on e-Learning (July 2009)
11. IAB, Cifras del Mercado de Internet en Chile. Informe preparado por Interactive Advertising Bureau Chile (2013), `http://www.iab.cl/estudios-iab/` (revisado March 13, 2014)
12. Inalef, K., Guerra, J., Scheihing, E.: Development and Validation of a Virtual Worklog as a Collaboration Tool for the Kelluwen Learning Community. In: 2012 IEEE 11th International Conference on IEEE Trust, Security and Privacy in Computing and Communications (TrustCom) (2012)
13. Paul, M.J., Dredze, M.: You are what you Tweet: Analyzing Twitter for public health. In: ICWSM (2011)
14. Scheihing, E., Guerra, J., Cárcamo, L., et al.: La experiencia Kelluwen: Tres años de desarrollo y puesta en práctica de una propuesta de innovación didáctica con uso de TIC. Estud. Pedagóg 39, 121–141 (2013)
15. Scheihing, E., Guerra, J., Bustamante, S.: An experience of peer evaluation in a b-learning environment. Paper accepted in EdMedia 2014 (2014)
16. Wright, N.: Twittering in teacher education: reecting on practicum experiences. Open Learning 25(3), 259–265 (2010), doi:10.1080/02680513.2010.512102
17. Yang, X., Ghoting, A., Ruan, Y., Parthasarathy, S.: A framework for summarizing and analyzing twitter feeds. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 370–378. ACM (August 2012)

## A Table of Didactic Designs

At each Principal Component Analysis, points are represented by the number of the Didactic Design instead of their Experience number. Note that some names are repeated, as they are newer versions of other the Design.

| Design id | Name (translated) |
|---|---|
| 21 | "The literary and non-literary in Youtube" |
| 23 | "Panoramic of our world through photograph" |
| 24 | "Newsblog" |
| 25 | "Greco-latin culture from XXI century blog" |
| 26 | "Photographing the biosphere" |
| 27 | "Touristic blog of our locality" |
| 28 | "Video run: The worlds in literature" |
| 29 | "Photographing the Industrial Revolution" |
| 43 | "Rebuilding our cultural heritage" |
| 65 | "Photographing the biosphere" |
| 66 | "Creating a Monarchy through a Blog" |
| 67 | "I tell my tale" |
| 68 | "A letter has arrived!" |
| 73 | "Our magazine of Chilean History" |
| 77 | "Blogging opinions, building realities" |
| 78 | "Speech of comments and its possibilities" |
| 80 | "Touristic blog of our locality" |
| 81 | "Photographing the Industrial Revolution" |
| 82 | "The literary and non-literary in Youtube" |
| 83 | "Video run: The worlds in literature" |
| 84 | "Greco-latin culture from XXI century blog" |
| 87 | "Newsblog" |
| 88 | "Building a Slide-show of the XX Century" |
| 102 | "Research blog about Conspiracies" |

# dmTEA: Mobile Learning to Aid in the Diagnosis of Autism Spectrum Disorders

David Cabielles-Hernández[1], Juan Ramón Pérez-Pérez[1], MPuerto Paule-Ruiz[1], Víctor M. Álvarez-García[2], and Samuel Fernández-Fernández[3]

[1] Departament of Computer Science, University of Oviedo, Spain
{UO209910,jrpp,paule}@uniovi.es
[2] Katholieke Universiteit Leuven, Belgium
victor.alvarez@cs.kuleuven.be
[3] Departament of Education Science, University of Oviedo, Spain
samuel@uniovi.es

**Abstract.** Mobile Learning is a teaching-learning methodology which has been developed successfully both inside and outside the classroom. Thanks to the several possibilities of mobile devices, it has been possible to use them with students with special educational needs thus giving rise to software applications focussed on competence acquisition on the students' part and offering such a high level of interaction that it could not be possible by using a PC. Advances in the development of mobile devices have made it feasible to go a step forward and help the teacher with the modelling and evaluation of this type of students. dmTEA is a mobile technology which allows both behaviour evaluation and modelling of students on the autism spectrum. It implements 12 activities – duly verified by experts – which are adapted to interaction with the mobile device within a specific context, the aim of which is to deal with student's disorders by observing learning processes and modelling his or her behaviour during task performance. dmTEA implementation in two studies, i.e. a student with severe autism and another student with moderate autism, has helped the teacher at the time of evaluating by offering him or her the necessary information to carry out any possible intervention.

**Keywords:** mobile learning, users with special needs, autism.

## 1 Introduction

During the last years, smartphones have undergone a great development both regarding hardware –thanks to their built-in sensors which allow context aware applications [1] – as well as software, which offers the possibility to explore these devices to their highest level to develop such applications. Besides, interest has been greater due to reduction of cost as well as to the support offered to users in different fields such as education [2–4].

Within the educational field, applications for Android and iOS mobile devices have given rise to innovative educational scenarios from pre-school education to high

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 29–41, 2014.
© Springer International Publishing Switzerland 2014

school education, based on access from everywhere and at any time. Context information, obtained from sensors available in the device, has given rise to the implementation of mobile learning [5, 6] in educational projects involving exhibitions, museums and topic parks. Besides, mobile learning is having a key role as a very useful tool for being used in the classroom.

"Educational Integration" is a process by means of which ordinary schools are looking for and producing the supports required by students with learning difficulties, special educational needs or with any type of disability. The expression “special educational needs” was first used in the 70s, but it was widely spread in the 80s by the Warnock Report, prepared by the Secretary of Education of the United Kingdom in 1978. In the specific case of Spain, the law on Education which is now in force, the LOE 2/2006 passed on the 3rd. of May, in its Title II, mentions the term ACNEAE (students with specific needs of educational support), as it refers to “any student who requires, during his or her complete schooling, or any part thereof, certain specific educational support and care as a consequence of disability or severe behaviour disorders” (section 73).

Among the students with special educational needs we may mention, among others, those belonging to the autist spectrum [7]. For autism diagnosis, teachers rely on specialists’ report and among other questionnaires, on the IDEA inventory (Inventory of the Autist Spectrum) [8], the main purpose of which is to assess the severity of autist features and how deep they are in a person who is older than 5-6 years old. IDEA has been chosen – out of other similar tests – because the other tests involve rating scales with a few items giving marks which are not highly relevant for the teacher. Besides, IDEA is a test with a wide theoretical and technical basis; it is more qualitative than others and thus, the teacher may analyze, in detail, the person’s potential and he or she may evaluate twelve characteristic dimensions of students with autist spectrum and/or serious development disorders. At the same time, dimensions are divided into four characteristic levels. The main aims of the IDEA inventory are to establish, during the diagnosis procedure, the severeness of the user’s autist features, to help the teacher to prepare learning strategies and to evaluate medium and long-term changes taking place in the student’s behaviour.

It is within this educational context where the dmTEA technology is developed. It gives the possibility to carry out, inside the classroom, an evaluation of users with autist spectrum, and it may be adapted to the specific needs by means of different activities. In this respect, dmTEA offers the teacher a detailled report with marks for each dimension. With the information obtained, families and experts have the possibility to plan an intervention together, focusing on the dimensions that exhibit a greater affectation status and improving them gradually with the most appropriate activities for each case based on the report data.

For this purpose, this study pursues two main objectives. The first one is to define and develop context-aware tasks which arise student’s interest and motivation, so that they may provide the basis for evaluating the disorders defined in IDEA. The second one is to gradually solve the possible disorders suffered by the student by modelling his or her behaviour during the performance of the tasks in class.

## 2 Background

The principles of “The Universal Design for Learning” (UDL) [9] encourage the offer of universal access to educational curricula for all students, thus ensuring equal opportunities. Moreover, in the particular case of education, The Universal Declaration of Human Rights (1948) (http://www.un.org/en/documents/udhr/) states the right to education and equal access to education for all on the basis of merit (Art 26).

The scientific community has been conscious of this need. Within this context, students with autist spectrum have been the focus of part of the development of educational systems using the benefits offered by mobile devices. The reason is that they offer several possibilities both in the expressive as well as in the receptive field, thanks to the reduction of linguistic contents in favour of graphism, iconicism and the multi-sensorial supports available. There is a wide range of applications which vary from more communicative aspects to others which are merely educational. Within the applications facilitating communication, we may mention some examples such as the builders of phrases based on pictograms - designed at the beginning of the year 2000 in PDAs [10] - to the most recent ones developed in mobile systems (DiegoSays; https://play.google.com/store/apps/details?id=com.benitez.DiegoDice). Talking about present educational applications, an important improvement in quality has been achieved with the use of mobile devices, specially those allowing the teacher to prepare the activities, in such a way so as to be able to increase the possible options both regarding activities – thanks to sensors - as well as in the search of acquisition of specific knowledge, specially focussing on the acquision of skills such as language, mathematics, awareness of environment, capacities regarding autonomous performance of activities or other social skills [11]. We may also have applications which teach organizational aspects by means of pictographic diaries, which offer simple exercises adapted to different subjects such as Language or Mathematics, which have the possibility to be adapted to the use of ordinary computer applications or which help in order to understand feelings by means of exercises using sensors [12].

From a pedagogical point of view, there are different personal computer applications which partially cover the dimensions offered by IDEA. For our work, we have made a selection of them (refer to Table 1), chosing those which give the possibility to cover the greatest part of said dimensions, focussing on the activities specifically related with autism patterns. In general, most of the applications have the following characteristics [13]:

- They offer a controllable situation and environment; they are a highly-predictable partner who offers perfect and understandable contingencies: by pressing the same key, the same results are always obtained.
- They present a multisensorial – mainly visual - stimulation, offering benefis to people with ASD (Autist Spectrum Disorder).
- Their effort and motivation capacity is really high, fostering attention and reducing frustration as a consequence of mistakes.

- They foster or give the possibility to carry out autonomous work and to develop self-control capacities. The TICs are adapted to personal characteristics, thus allowing different learning rates and a greater level of individualization.
- They are an active learning element with versatility, flexibility and adaptation as the main features.

**Table 1.** PC applications for autist spectrum disorders

| Application | Aim | Author/Publisher/Web |
|---|---|---|
| "Sócrates 102 activities" | It allows the design of tasks aimed at distinguishing figures, colours and geometrical shapes which allow the evaluation of mental flexibility and receptive language features. | EMME Interactive |
| "Adibú" | It has exercises which guide the tasks towards receptive language disorders, anticipation and suspension, by means of exercises focussed on solving problems and ocular and motor coordination. | COKTEL EDUCATIVE |
| "Responsive Face" | It allows to create stories and watch animated films, fostering the recognition of face expresión and feelings. | http://www.mrl.nyu.edu/~perlin/facedemo/ |
| "Clic" | It gives the possibility to relate images and to orally interact by means of questions and replies or by learning new nouns and adjectives. | Francesc Busquets http://www.xtec.es/recursos/clic |

## 3 dmTEA: Mobile Tool for the Diagnosis Evaluation of Autism Spectrum Disorders (ASD)

The previously-described background is the basis for designing an evaluation-diagnosis tool for autism which includes a set of learning activities based on the IDEA inventory using, for this purpose, the technology of mobile devices called "dmTEA". Specifically, IDEA estimates autism by means of 12 dimensions which are the main disorders that define it. These dimensions are gathered creating four blocks which correspond to the four sections mentioned by Lorna Wing [7]: Socialization, language and communication, anticipation and flexibility, and symbolization (Table 2).

**Table 2.** IDEA dimensions regarding the main disorders

| **Disorders** | **Dimensions** | | |
|---|---|---|---|
| ***Socialization*** | 1. Social relationship | 2. Joint reference | 3. Inter-subjective and mentalist |
| ***Communication and language*** | 4. Communicative functions | 5. Expressive language | 6. Receptive language |
| ***Anticipation and flexibility*** | 7. Anticipation | 8. Flexibility | 9. Meaning of the activity |
| ***Simbolization*** | 10. Fiction | 11. Immitation | 12. Suspension |

The final aim of the activities designed is to work on disorders in order to achieve, by behaviour modelling, gradual solutions at the same time that students acquire the necessary knowledge and competentes. From a number of 26 tasks initially suggested, experts – in this case, the teacher and two professional experts on autist disorders from our University– have chosen 12. The criterion is based on choosing the most suitable tasks for interacting with the mobile device for the specific context. The learning activities designed are (Fig. 1):

1. Interaction between the child and an adult by asking the child to press on a specific figure, which will be a square or a circle. To choose the figure, press his or her finger on the touchscreen.
2. On the screen, drag a ball towards a child. For this purpose, the student presses on the ball with his or her finger and drags it on the screen towards the child's image.
3. To imitate applauses and greetings seen in a video on the screen. The student will imitate the gestures of the video and appraisal will be manual by pressing on a button on the screen.
4. To take objects to their profiles, which are symbolised by pictograms, in such a way that the student should place them in order and he or she must take them, one by one, to their corresponding mould. For this purpose, he or she must press his or her finger on the object and drag it along the touchscreen to its outlined shape.
5. To establish the sequencial steps to go to school by means of several pictograms following a previously-defined order. There will be some boxes into which the student must place the pictograms. For this purpose, each pictogram will be chosen by pressing his or her finger on it and dragging it towards the chosen box.
6. To make an agenda with pictograms for one day. In this case, the order may be the one chosen by the student as the suitable one. As in task 5, the student will have some boxes for the pictograms to be dragged into them with his or her finger.

7. To forecast the weather based on a picture on the screen of a man holding an umbrella. With his or her finger, the student will press on a pictogram which reproduces the specific weather situation.
8. To point out the different types of mood requested, which are represented by different icons which simulate feelings. For this purpose, the student will press his or her finger on the icon corresponding to the feeling requested.
9. To differentiate and learn adjectives choosing the one requested by the task out of the two options shown on the screen, such as open-closed, big-small, etc. The student will press on the touchscreen choosing the picture which represents the adjective requested.
10. To repeat the name of the objects, as the mobile device reproduces the names when the student presses on their pictures on the screen. Once the student listens to the name, he or she repeats it. Such repetition is heard thanks to the microphone and the software of the mobile device interprets it to establish if it is valid or not.
11. To decide when the student may cross the street based on the traffic lights appearing on the screen and which will change from red to green after a short period of time. For this purpose, the student presses his or her finger on the pedestrian appearing on the picture and he will move if the traffic lights are green or he says crossing is forbidden because the traffic lights are red.
12. To paint a tree on a screen picture with different colours and thicknesses. On the right, the student will find two colours, i.e. greean and brown, and two thicknesses. These elements may be chosen by pressing on them. To paint, it is only necessary to drag his or her finger on the screen after having chosen the colour.

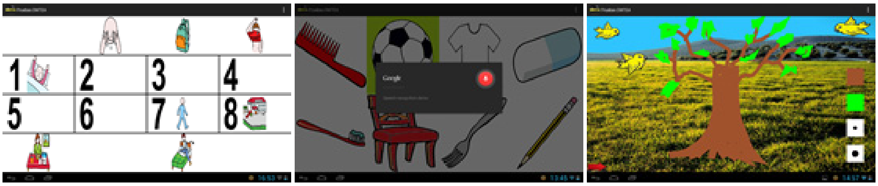



**Fig. 1.** Screenshots corresponding to the picture agenda tasks (5). Repeat the object names (10) and paint a tree (12).

In the following step, experts decide the specific aspects of IDEA (Table 3) which are most suitable for being evaluated with electronic devices. Selection is based on deciding which aspects could be more related with the tasks chosen in the previous step, in order to optimize the result of the tests, taking into account a maximum potential relation between dimension and task, and rejecting the most complicated ones or those which are impossible to be assessed with a mobile device. This is because IDEA shows dimensions, such as those of social relationship, which are nearly impossible to be estimated automatically by using a device because real interaction with another user must be checked in person.

**Table 3.** Dimensions assessed in the evaluation

| Dimensions | Description |
|---|---|
| Dimension 1 | Non-frequent, induced and external relations with peers.<br>Relations are more a response than the consequence of own initiative |
| Dimension 2 | Use of joint reference looks in directed but non-open situations.<br>Established guidelines of joint action and attention but not joint worry. |
| Dimension 5 | Language made up of loose words or echolalias. There is no formal creation of phrases and sentences.<br>Sentence language. There are some sentences which are not echolalic, but they do not create speech or conversation. |
| Dimension 6 | Literal and nearly inflexible understanding of the statements, with some kind of structural analysis. The speech cannot be understood.<br>Conversation is understood but the difference between literal and intentional meaning could hardly be distinguished. |
| Dimension 7 | Simple anticipatory behaviour in daily routines. Frequently, opposition to changes and worsening in situations which imply changes. |
| Dimension 8 | Complex rituals. Excesive affection towards objects. Obsesive questions.<br>Obsesive and limited contents of thought. Nearly unfunctional and flexible interests. Strict perfectionism. |
| Dimension 10 | Symbolic game which is generally not really spontaneous and obsessive. Important difficulties to distinguish between fiction and real world. |
| Dimension 11 | Simple, recalled and non-spontaneous motor imitations.<br>Established imitation. Lack of inside patterns. |

## 4 Methodology

With the taks and dimensions chosen, we have designed an assay with which the tests are going to be carried out with the users (Fig. 2). This assay is a simplified version of the Delphi Method [14], in which other experts make assessments of tasks and dimensions, filling in a matrix form in which the rows are the dimensions while the columns represent the tasks. Specialists give a mark from 0 to 10 regarding the relationship between the taks and the dimension, being 0 an invalid relation and 10 an excellent result in the task.

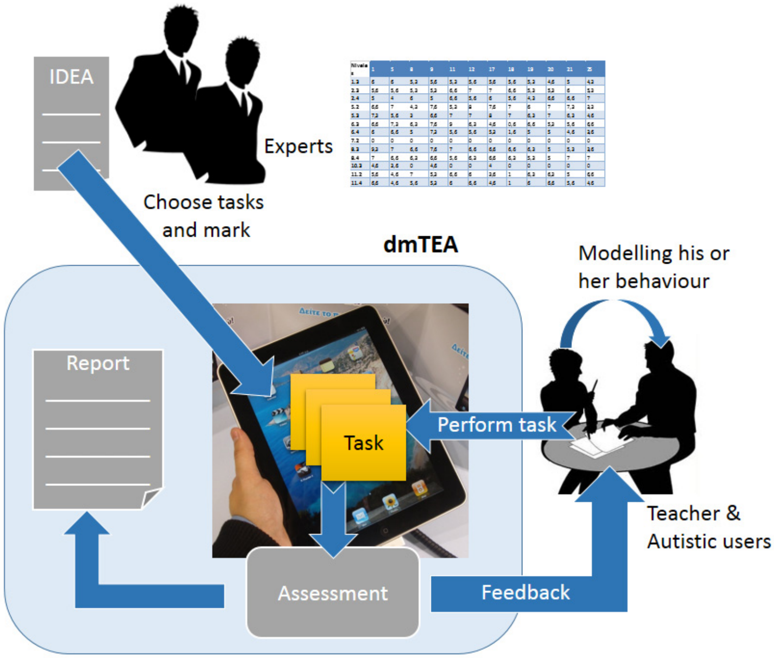



**Fig. 2.** Diagram which represents the use and assessment of dmTEA

Tasks of the assay are carried out with two users, one suffering from severe autism (Kanner type) and the other one with moderate autism (High-Functioning type) of the ADANSI centre (Association for the Support of Families with Autistic Members; http://www.adansi.es/). To perform the tasks, we have been helped by the teacher of the centre with the co-operation of 3 experts, who observe the performance of the tasks by each user. Based on this observation, each party will give the mark considered as the most suitable one based on his or her experience. Assessment is not given based on the proper performance of the task as the user's behaviour while performing the task is the clue for assessing the level of the user's disorder. Therefore, an exercise may be left incomplete but it will not necessarily mean that the user has some features of such disorder.

**Fig. 3.** Test of a task with the teacher and real users

All the tests are carried out in the institution to avoid the students' lack of concentration. The teacher will be in charge of giving the main orders, as they are the persons with whom they have a closest contact. In this way, it is avoided to have problems getting the expert confused when giving marks and that a mistaken assessment of the test is achieved because it is considered that it is not suitable, when the real problem is the lack of trust on other persons. All tasks have a feedback, which arrises the student's interest and he or she will be congratulated if it is properly performed or stimulated to do it again if he or she has failed. Besides, all activites may be repeated modelling the student's behaviour, thus allowing a better understanding of what it is requested in the task and being able to perform it properly by following the teacher's instructions.

These tests are carried out during two weeks, and two tasks are performed daily in order not to make the student feel tired. The first week is focus on sessions during which the students are in contact with the tasks while the second week is focussed on consolidation in order to understand the tasks in a better way and to perform them efficiently.

## 5 Results and Discussion

The aims of this analysis are to determine, by means of observation, the applicability and compliance of the tasks with the IDEA's pattern and to check the agreement between observers at the time of making the assessment in the matrix form.

**Table 4.** Final averages for moderate autism

| | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 | T12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D1 | 6.0 | 6.0 | 5.3 | 5.6 | 5.3 | 5.6 | 5.6 | 5.6 | 5.3 | 4.6 | 5.0 | 4.3 |
| D2 | 5.3 | 4.8 | 5.7 | 5.2 | 6.6 | 6.3 | 6.5 | 6.1 | 4.8 | 6.0 | 6.3 | 6.2 |
| D5 | 7.0 | 6.3 | 3.7 | 7.1 | 6.2 | 7.5 | 7.8 | 7.0 | 6.2 | 7.0 | 6.8 | 4.0 |
| D6 | 6.3 | 7.0 | 5.7 | 7.5 | 7.3 | 6.0 | 5.0 | 1.1 | 5.8 | 5.2 | 5.1 | 5.1 |
| D7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| D8 | 8.2 | 6.8 | 6.5 | 7.1 | 6.3 | 6.5 | 6.6 | 6.5 | 5.8 | 5.0 | 6.2 | 5.3 |
| D10 | 4.6 | 3.6 | 0.0 | 4.6 | 0.0 | 0.0 | 4.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| D11 | 6.1 | 4.6 | 6.3 | 5.3 | 6.3 | 6.3 | 4.1 | 1.0 | 6.2 | 6.5 | 5.3 | 5.6 |
| Berk index | 0.9 | 0.8 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 1 | 0.9 | 0.9 | 0.9 | 0.8 |

**Table 5.** Final averages for severe autism

| | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 | T12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D1 | 8.3 | 8.3 | 6.0 | 7.3 | 9.6 | 6.0 | 7.3 | 6.3 | 6.3 | 7.0 | 8.0 | 6.3 |
| D2 | 5.3 | 1.8 | 5.7 | 2.0 | 9.2 | 5.5 | 4.5 | 6.7 | 5.8 | 5.5 | 3.0 | 5.5 |
| D5 | 0.0 | 0.0 | 6.2 | 0.0 | 6.2 | 0.0 | 7.0 | 5.5 | 6.0 | 6.0 | 3.2 | 7.2 |
| D6 | 9.0 | 8.0 | 5.5 | 10.0 | 8.1 | 2.5 | 4.3 | 5.0 | 3.8 | 7.2 | 8.3 | 8.2 |
| D7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| D8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 4.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| D10 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| D11 | 7.8 | 8.0 | 8.0 | 7.7 | 2.3 | 6.5 | 7.2 | 7.2 | 5.7 | 7.0 | 6.2 | 0.0 |
| Berk Index | 1 | 1 | 0.8 | 0.9 | 1 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 1 |

The value of the Berk index (agreement between observers), shows that in both cases, there is an excellent level of agreement between experts for all the tasks. Based on the results obtained, the average rates of the marks collected by the observers may be estimated as well as the average correlation level of task and dimension (Tables 4 and 5).

In the case of severe autism (Table 5), it can be observed that the most saturated dimensions due to the tasks carried out are 1 (social relationship), 6 (receptive language) and 11 (imitation). Thus, it can be declared that the set of tasks gives us the possibility to establish first estimations of the functional diagnosis about such dimensions. It may also be observed that relationships are not bi-univocal, i.e., tasks are not related with all the dimensions in the same way. The most consistent relations may be observed in task 5 (to establish the steps to go to school) with dimensiones 1 (social relationship) and 2 (capacities of joint reference), in task 4 (to put objects to their outline shapes) with dimension 6 (receptive language) and in task 2 (to move a ball towards a child) with dimension 11 (imitation).

Results shown in Table 5 show the student's communication problems. The teacher informs that communication relations between the student and himself or herself are induced by the latter, which means that conversation only takes place when the teacher encourages it. Therefore, we may declare that, for this specific case, the information shown in the Table supports the aims initially specified in this analysis of applicability and adjustment of the tasks to the IDEA paradigm. Thanks to such aim achivement, dmTEA allows the teacher to detect, empirically, the IDEA dimensions in which the student commits mistakes, supplying information for a possible intervention.

In the case of moderate autism (Table 4) we have that the highly-saturated dimensions are 5 (expressive language), 6 (receptive language) and 8 (mental and behavioural flexibility). As in the previous case, this information gives the possibility to make initial assessments of functional diagnosis with reference to the dimensions mentioned. In this case, the most consistent relations have taken place in task 8 (specify moods) with dimension 1 (social relationship), in task 4 (to put objects to their outline shapes) with dimension 6 (receptive language) and in task 7 (weather forecast) with dimension 5 (expressive language).

Results of Table 4 (moderate autism) show the problems regarding behaviour and language the student has, both of expressive and receptive nature. The teacher informs that at the time of performing the tasks, the student is anxious, which has gradually been solved with the teacher's help as he or she is in charge of calming down the student while explaining orally and modelling the test performance. Therefore, we may confirm, as in the previous case, that information included in the table supports the aims initially laid down in this analysis of applicability and adjustment of tasks to the IDEA paradigm. For this second case, as in the previous one, dmTEA offers the teacher the necessary information – based on the marks for each dimension – to carry out a possible intervention apart from being the support for diagnosis evaluation.

Apart from the data obtained regarding the relationship of tasks and dimensions, observation gives us the possibility to confirm that it is possible to shape the user's behaviour. Finally, the student understands the task and performs it properly; it is not solved by the trial-and-error method and thus, it is possible to solve the most evident disorders little by little. The teacher and experts also declare that users suffering from behavioural and emocional disorders receive a clear benefit when interacting with the tablet, something which would be practically impossible using a mouse or a keyboard.

## 6 Conclusions

Integration of mobile devices in education has given rise to a teaching and learning methodology called Mobile Learning. Characteristics such as access from anywhere and at any time give the possibility to create contextual educational scenarios, producing educational applications which may be used both inside and outside the classroom by teachers and students.

In this paper, we are showing a different scenario of mobile learning application and specifically, its use and applicability regarding the evaluation and modelling of behaviour, inside the classroom, of special educational needs, and specifically, for students with problems regarding behaviour and language. Up to now, mobile devices have been used within the autism disorder field, with the aim of making students acquire certain competences but not as an element which may help and support teachers for evaluating and modelling students in their classrooms.

dmTEA is a software technology which permits the evaluation of autist spectrum. Said evaluation is based on the IDEA inventory, which evaluates twelve characteristic dimensions of students with autist spectrum and /or with deep development disorders. This technology implements 12 learning activities adapted to interaction with the mobile device for a specific context and its aim is to work with disorders to solve them gradually by modelling behaviour and, at the same time, students acquire the necessary knowledge and competences.

The experimental use of the dmTEA in two cases, makes us declare that it is a tool which helps at the time of evaluating the educational possibilities of this type of students apart from offering the teacher the necessary information – based on the marks for each dimension – to carry out a possible intervention. The design of the tasks in dmTEA benefits the modelling of students' behaviour, which finally results in the understading of the task and its proper performance by the students.

Results obtained encourage furthering this study. Directions of future research include both extending the number of users involved in the evaluation as well as including the four existing levels of autism (Kanner, Regressive, High-Functioning and Asperger), in order to be able to generalize the observed behaviour to situations that were not examined. Other future lines of work are: complete dmTEA with all the IDEA dimensions and include more tasks; further development the diagnosis system to facilitate the present management of the experts' reports and create new activities which can progressively help with disorders; and complement dmTEA with other evaluation instruments for girls and boys from 2 years on, such as CARS (Childhood Autism Rating Scale) or ADOS (Autism Diagnostic Observation Schedule) which, besides autism diagnosis, open the possibility of using dmTEA for other type of special educational needs.

**Acknowledgments.** This work has been funded by the Department of Science and Innovation (Spain) under the National Program for Research, Development and Innovation: project TIN2011-25978 entitled "Obtaining Adaptable, Robust and Efficient Software by including Structural Reflection to Statically Typed Programming Languages".

## References

1. Afyouni, I., Ilarri, S., Ray, C., Claramunt, C.: Context-aware modelling of continuous location-dependent queries in indoor environments. J. Ambient Intell. Smart Environ. 5, 65–88 (2013)

2. Sharples, M.: The design of personal mobile technologies for lifelong learning. Comput. Educ. 34, 177–193 (2000)
3. Elias, T.: 71. Universal instructional design principles for mobile learning. Int. Rev. Res. Open Distance Learn. 12, 143–156 (2011)
4. Park, Y.: A pedagogical framework for mobile learning: Categorizing educational applications of mobile technologies into four types. Int. Rev. Res. Open Distance Learn. 12, 78–102 (2011)
5. Traxler, J., Kukulska-Julme, A.: Mobile Learning in Developing Countries. Commonwealth of Learning (2005)
6. Ally, M.: Mobile learning: Transforming the delivery of education and training. Athabasca University Press (2009)
7. Wing, L., Gould, J.: Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. J. Autism Dev. Disord. 9, 11–29 (1979)
8. Rivière, Á., Martos, J.: Tratamiento del autismo. Nuevas perspectivas. Instituto de Migraciones y Servicios Sociales, Madrid (1998)
9. Rose, D.H., Meyer, A.: Teaching every student in the digital age: Universal design for learning. ERIC (2002)
10. Gea-Megías, M., Medina-Medina, N., Rodríguez-Almendros, M.L., Rodríguez-Fórtiz, M.J.: Sc@ut: Platform for Communication in Ubiquitous and Adaptive Environments Applied for Children with Autism. In: Stary, C., Stephanidis, C. (eds.) UI4ALL 2004. LNCS, vol. 3196, pp. 50–67. Springer, Heidelberg (2004)
11. Fernández-López, Á., Rodríguez-Fórtiz, M.J., Rodríguez-Almendros, M.L., Martínez-Segura, M.J.: Mobile learning technology based on iOS devices to sup-port students with special education needs. Comput. Educ. 61, 77–90 (2013)
12. Leijdekkers, P., Gay, V., Wong, F.: CaptureMyEmotion: A mobile app to improve emotion learning for autistic children using sensors. In: Proceedings of CBMS 2013 - 26th IEEE International Symposium on Computer-Based Medical Systems, pp. 381–384 (2013)
13. Nicolás, F.T.: Tecnologías de ayuda en personas con trastornos del espectro autis-ta: guía para docentes. CPR Murcia I (2004)
14. Linstone, H.A., Turoff, M.: The Delphi method: Techniques and applications (1975)

# Use of Social Media across Different Generations in Higher Education in a Developing Country

Ilknur Celik[1] and Cyril Schoreels

[1] Middle East Technical University Northern Cyprus Campus
cilknur@metu.edu.tr

**Abstract.** While social media is increasing its extent and reach every day, research shows that its impact seems to be more intense in developing countries for various reasons. In developing countries, social media not only satisfies personal communication needs but also tends to compete with mainstream media for news and play a significant role in social movements. Hence potential use of this mounting medium for education needs to be explored. In this study, a social media and education survey has been prepared in order to measure perceptions, awareness and concerns of students and educators coming from different generations on the topic of social media and its use in education. Such a study is crucial to materialize different policies and strategies for enhancing education with this popular and freely available technology. Students' and faculty members' perceptions, expectations and willingness about the educational use of social media are analyzed and different generations are compared. Results of this study could shed light for developing various strategies for integration of social media into Higher Education in developing countries where the difference between generations is typically felt larger.

**Keywords:** Social Media, Technology Enhanced Learning, Generation Divide, Survey Study, Higher Education, Education Policies, Developing Countries.

## 1 Introduction

Social media has received a widespread adoption over the last decade worldwide for various purposes from communication to exchange of ideas, marketing, networking, following news or celebrities, and even utilized as a tool for social movements. This rising uptake of social media is not falling back much in developing countries neither even though one might think that the infrastructure, slower adoption of technology, social, economical and cultural aspects may hinder its popularity. In our case, we observe that social media is playing a very important role in people's everyday lives in Turkey (a developing country according to the country classification in the *World Economic Outlook* by International Monetary Fund[1]) and its popularity is growing day by day especially among the young people as well as the older generations.

---

[1] IMF World Economic Outlook Database -WEO Groups and Aggregates Information https://www.imf.org/external/pubs/ft/weo/2014/01/weodata/groups.htm, last accessed in April 2014

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 42–55, 2014.
© Springer International Publishing Switzerland 2014

Growing popularity of social media could be presumed for communication and social sharing in general. However, social media's role in higher education is still a common topic of debate. Although many educational institutions have quickly embraced social media for various purposes outside the classroom, the extensive if and how discussions have been acting as an impediment for its adoption into formal education. Hence, many researchers worldwide are studying how social media can be integrated into classes, or whether or not it can be implemented to enhance learning, collaboration, information sharing and community building in higher education.

A number of interdisciplinary researchers studied the effects of social media and education amalgamation on learning and teaching as well as the limitations and concerns from pedagogical point of view. There are many case studies where some form of social media tools and technologies were experimented with, for instance, in a specific class [1], subject areas [2], and training in a specific field of students [3]. Tess [4] and Davis III et al. [5] provide an overview of a number of studies carried out regarding the use of social media in higher education. There are also studies showing the attitudes of faculty members towards specific social media platforms. For instance in a study by Roblyer et al., who tried to measure the social networking platform Facebook's future in education, found that Facebook was mostly perceived as social rather than for education, but concluded that the students were more open to the idea of its use in education than the faculty members [6]. Another study in the UK also showed that students identify Facebook as social rather than as a tool for formal learning, although they sometimes employ it informally for learning purposes such as collaboration, but were not keen on being contacted by tutors via this platform [7]. There are also studies that examine students' perceptions of faculty members on various social media services. For instance, a study examined the impact of posting social, scholarly, and a combination of social and scholarly tweets to the social networking site Twitter and their impact on perceived teacher credibility [8]. The results of this research showed that the students who saw only the social tweets rated the faculty member higher in perceived credibility than the group that saw only the scholarly tweets, suggesting that instructors should embrace these new opportunities to disclose information about themselves on social networking sites, as this could lead to increased perceived credibility and in turn, may lead to increased student motivation, a greater interest in the material presented in the classroom, and a greater willingness to learn from the instructor.

Literature tells us that the adoption of social media could be beneficial for education and a step up in the technology enhanced learning paradigm. The state-of-the-art today highlights the need to make these new modes of digital conversation and discourse become more mature, efficient, safe, and truly useful for education, and fortunately much can be done in higher education in order to initiate the change. While there are numerous studies which tried to determine the effects of social media in education from students' and some from faculty members' standpoint, it is important to study the beliefs and attitudes of both parties in order to determine how such adoption could take place as well as their readiness or acceptance for this.

Firstly, we believe that the differences in opinions or attitudes towards social media are more likely to come from different generations rather than different professions. Furthermore, we consider that social, cultural and economical conditions may have effects on such adoption when developing countries are compared to developed countries.

Generations are defined as “time intervals which are formed with worldwide economic and social movements” [9]. Due to changes in their upbringing styles and environments where they live in, important differences can be observed between generations in terms of their characters, expectations and working methods [10]. Generation categories are defined as “Baby Boomers” for those born between 1943 and 1960, “Generation X” between 1961 and 1981, “Generation Y” or “Millennials” between 1982 and 2004, and “Generation Z” or “Homelanders” for those born after 2005[2].

Social generations are cohorts of people who were born in the same date range and share similar “cultural experiences” and this may differ with countries in different geographic locations. According to a comparison study of the identified characteristics of Generation Y in Europe and USA with Turkey, for example, found that most of these key characteristics attributed to this generation were common to all these geographic locations although some habits, behaviors and attributes like online shopping, patience, career planning, soft, technical and technological skills differed slightly[11]. Generation gap might be felt stronger in developing countries than developed ones due to belated access to technology, slower adoption, and so on. This might mean that we may have a stronger difference between the “Baby Boomers”, the “Generation X” and the “Generation Y” in developing countries. In today’s world where different age groups work together, an important part of the problems may stem from perception, method, practice and communication differences. These issues have been widely studied in education research community [12], [13], [14].

In view of that we conducted a survey research to measure the social media usage, attitudes, beliefs and experiences of undergraduate students and faculty members in a top reputable university in Turkey. Since this institution has been a pioneer in introducing various technologies to the country and the people in the past, it was considered as the best choice to investigate: *1- how open the students and instructors are to the idea of using social media*; *2-their awareness of educational use of social media*; and *3- if and how these differ with different generations*. By examining the situation in this higher education institution, we might infer certain policies for educational use of social media in other higher education institutions throughout the country.

## 2 Study Design

A survey research was conducted to measure the social media use and habits of students and faculty members as well as their perceptions and attitudes towards social media in general and in education. A cross-sectional study was chosen to collect data in a short amount of time in order to eliminate time factor affecting the results, as well as to reach widest participation across the university. An online questionnaire method

[2] Generations in History: `http://www.lifecourse.com/about/method/timelines/generations.html`, last accessed in April 2014.

was selected as the most suitable survey instrument to study the viewpoint of the students and faculty members regarding social media use.

The questionnaire was developed both in line with related work done in the field and according to the interview responses that were obtained from a study group. For this reason, a series of quantitative survey interviews were implemented with a small group of volunteer faculty members and students in order to develop the questionnaire. These participants were asked semi-structured interview questions which provided response options to interviewees. These responses were recorded and analyzed in order to determine and design the data collection instrument. Special attention was paid to avoid double-barreled, leading, ambiguous or double negative questions.

The resulting questionnaire consisted of a **"Demographics"** section which contained questions about age and gender; a "**Social Media Usage Habits**" part to measure general, day-to-day social media habits; and a "**Social Media and Education**" part to measure perceptions on the use of social media in education including motivational and emotional questions. The latter two sections of the questionnaire contained a series of close-ended questions with a 4-point Likert scale (i.e.: *strongly agree*, *agree*, *disagree*, and *strongly disagree*).

In the **Social Media Usage Habits** section, participants were asked to state which, if any, of the social media platforms they use from a given list of different social media applications; declare how often and for what purposes they use these platforms for; and identify whether they are producers and/or consumers of social media by answering if they create certain digital content. Participants were then asked about their opinions, beliefs, concerns and attitudes regarding social media use in today's digital world based on their practices or perceptions.

In the **Social Media and Education** part, participants were first asked if they use social media platforms for educational purposes. For those who replied "yes", a variety of platforms were inquired to determine which ones were common and trendy. Participants were then presented with questionnaire items regarding their thoughts on using social media in education. These questions were designed to measure their concerns, motivations and attitudes toward social media use in higher education.

Special care was given not to include any other personal questions to keep the participants anonymous in order to maximize response rate. After reliability and validity tests were carried out and the required approval was obtained from the relevant ethics committee, a pilot study was carried out for the questionnaire items. In the light of the factor analysis, certain items were removed from the survey and/or rearranged as explained in detail in [15]. The modified survey was sent for expert opinion for feedback (with a 0.80 interrater reliability) and was reshaped into its final form as a web-based questionnaire.

Regarding the large population and high number of the variables, the stratified random sampling was not possible for the implementation of the questionnaire. For this reason, the questionnaire was published online via a web-based survey service (https://survey.metu.edu.tr/) in November 2013 and was accessible for participants for a period of two months. During this time, students and faculty members were contacted via university mailing lists to take part in the survey with periodic reminder emails to call for participation.

# 3 Results and Findings

The participants were undergraduate students as well as full-time and part-time faculty members from one of the most prestigious technical universities in Turkey. They were asked to complete the online questionnaire concerning their use and perception of social media anonymously.

The questionnaire was administered to a total of 1831 people, 1606 students and 225 faculty members (out of a potential population of ~<20000). Of 1606 students and 225 who responded to the questionnaire, only 1028 and 135 questionnaires were analyzed respectively as the others had insufficient data for analysis.

The analyzed responses of student participants (47.37% female, 52.63% male) showed a wide variety in terms of age (M=21.69, SD=3.7 years), and departmental affiliation (both from engineering & natural sciences (58.27%) and social sciences (41.73%)). The respondents were composed of preparatory school students (19.94%), freshmen (26.17%), sophomores (12.74%) juniors (15.66%) and seniors (14.59%). There were also respondents (10.89%) stating that they are in their fifth year of higher education who are the students that were enrolled in 5-year-long higher education programs offered in the university.

As for faculty members, the 135 responses (67% female, 33% male) also showed a wide variety in terms of their age (M=39.5, SD=9.2 years), departmental affiliation (both from engineering & natural sciences (33%) and social sciences (67%)) and professional experience (M=13.24, SD=9.2 years).

## 3.1 Generations

Since our hypothesis stated that generation gap might be a factor in people's perception, awareness and willingness to use social media platforms in their daily lives as well as for educational purpose, instead of comparing students' versus faculty members' responses, we regrouped all participants according to their generations and carried out data analysis accordingly.

The resulting dataset as a whole had a good distribution of age (mean=24, median=21) and gender (577 female and 586 male). The responses were split into the three previously defined generational categories as Generation Y (89.51%), Generation X (9.37%) and Baby Boomers (1.03%). Perhaps reflective of social media penetration in itself, we found a comparatively small sample for the Baby Boomers (BB), with more respondents falling into other categories. As expected, majority of the participants fell into Generation Y, as this category includes younger faculty members as well as the majority of student participants, whose numbers were considerably larger in the study, which is also in line with the distribution of the population in the sampling frame. According to Turkish Statistical Institute[3], 25% of the population of Turkey is Generation Y.

---

[3] Turkish Statistical Institute, `http://www.turkstat.gov.tr/`, last accessed in June 2014.

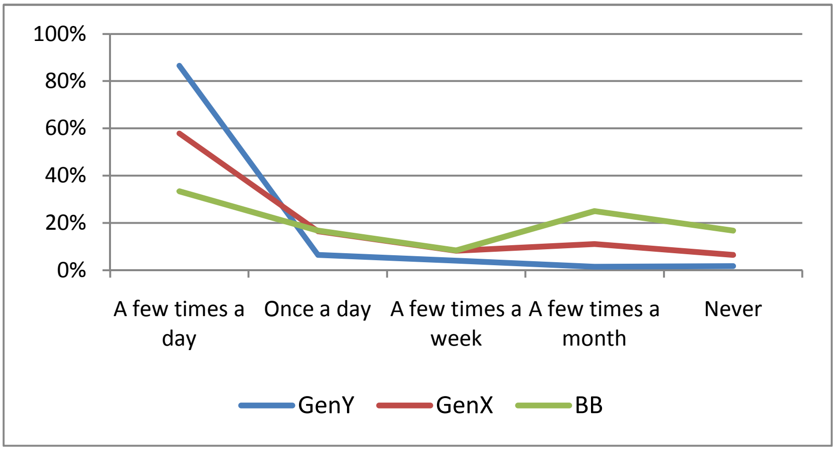



**Fig. 1.** General social media use frequency across generations

### 3.2 General Social Media Use

Firstly, we looked into general social media usage across generations such as ratios of usage, different platforms used and purposes for usage per platform. Data analysis shows that social media use appears most frequent amongst Generation Y users, with 87% making repeat uses during a single day. Generally, 93% of Generation Y access social media at least once a day, with only 74% by Generation X and a substantially lower 50% by BB as can be seen in Fig. 1.

Only a very small percentage of the Generation Y, namely 2%, stated that they do not use social media at all. Similarly, only 6% from Generation X declared the same while this goes up to 17% for BB. This indicates that social media might be the right choice of communication channel for students and lecturers if not now, in near future.

The participants were given a list of freely available and popular social media platforms and asked to indicate if and how often they use such services. Most respondents do not use the majority of platforms listed, irrespective of generational group, with a fairly consistent distribution across platforms across generations. Table 1 and Table 2 present a breakdown of frequency of use across selected social media platforms, for Generation X and Y.

**Table 1.** Frequency of use of selected social media platforms by Generation Y

| | *FB* | *Twitter* | *Google+* | *LinkedIn* | *Video* | *Picture* | *Bookmark* | *Blogs* |
|---|---|---|---|---|---|---|---|---|
| A few times a day | 81% | 36% | 11% | 2% | 56% | 17% | 6% | 8% |
| Once a day | 7% | 8% | 6% | 3% | 15% | 8% | 3% | 6% |
| A few times a week | 4% | 13% | 13% | 9% | 17% | 10% | 7% | 15% |
| A few times a month | 2% | 10% | 25% | 17% | 5% | 9% | 8% | 19% |
| Never/Not a member | 6% | 32% | 46% | 69% | 8% | 56% | 76% | 52% |

**Table 2.** Frequency of use of selected social media platforms by Generation X

| | *FB* | *Twitter* | *Google+* | *LinkedIn* | *Video* | *Picture* | *Bookmark* | *Blogs* |
|---|---|---|---|---|---|---|---|---|
| A few times a day | 53% | 22% | 10% | 4% | 16% | 5% | 3% | 3% |
| Once a day | 15% | 8% | 2% | 7% | 17% | 6% | 4% | 6% |
| A few times a week | 10% | 12% | 10% | 16% | 30% | 6% | 4% | 13% |
| A few times a month | 11% | 13% | 20% | 33% | 25% | 18% | 13% | 18% |
| Never/Not a member | 11% | 45% | 58% | 40% | 13% | 65% | 77% | 60% |

The major exception here is Facebook, which is the most popular amongst Generation Y users with frequent intraday access. This also extends to Generation X users, albeit to a marginally lower degree at 53% versus 81%. Another medium that is very popular only amongst the Generation Y users is video sharing platforms, with 56% accessing it repeatedly intraday and a cumulative 71% at least once a day. Generation X sees a fairly equal distribution in this social media type, and a low 16% repeat intraday accesses and a cumulative 33% at least once per day. Comparatively we saw no respondents using video sharing platforms regularly amongst BB.

### 3.3 Social Media Habits: Frequencies, Aims of Use, etc.

We investigate the primary motivation behind social media use by broadly dividing it into Personal and/or Academic, focusing on social media platforms such as Facebook, Twitter, Google+, LinkedIn, Blogs, Video sharing, Photo and Bookmark sharing.

The following analysis considers both broader uses in terms of the number of services employed for Personal and/or Academic purposes, as well as their generalization as shown in Fig. 2, Fig. 3 and Fig. 4. Overall, exclusively Personal use dominates the employment of social media across all three classes, with an excess of 2/3rds of all services not employed in the respondents' academic life.

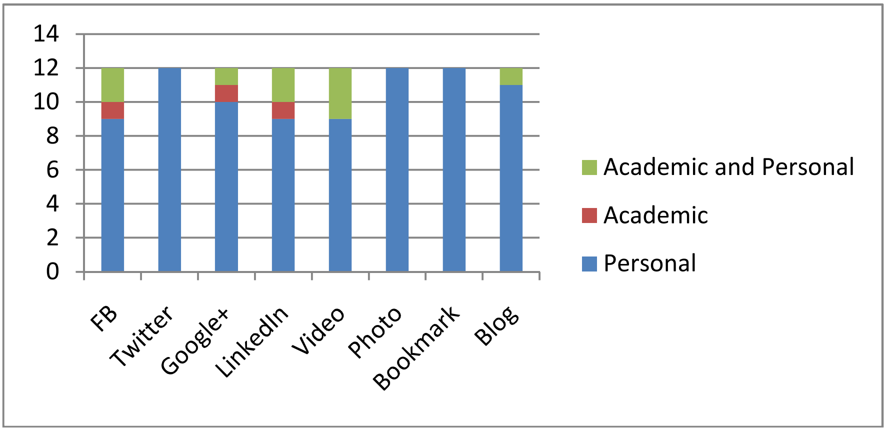



**Fig. 2.** Reasons for using selected social media platforms by Baby Boomers

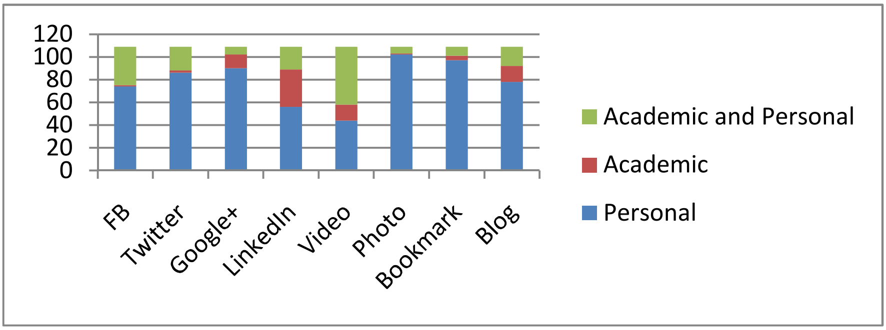



**Fig. 3.** Reasons for using selected social media platforms by Generation Y

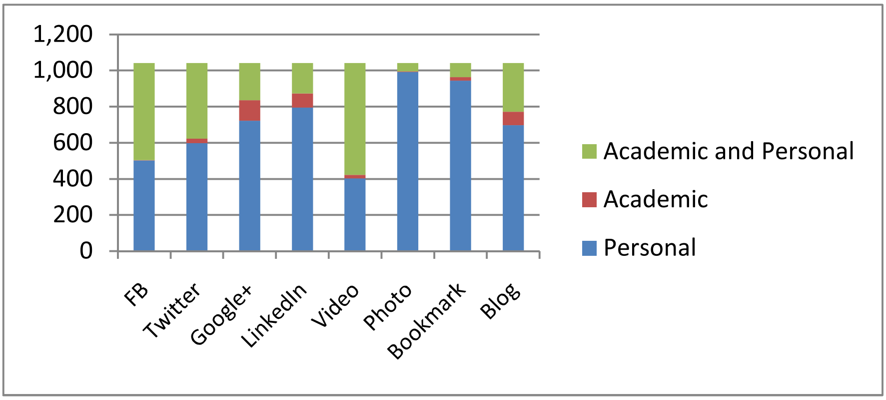



**Fig. 4.** Reasons for using selected social media platforms by Generation X

In the BB category (Fig. 2), personal use dominates, with no adoption of sharing mechanisms outside networking sites (i.e. Facebook, G+ and LinkedIn). Based on much larger sample sizes, Generation X and Y however form a stark contrast, with the majority using social media for both Academic and Personal use (demonstrated in Fig. 3 and Fig. 4). This suggests a much stronger penetration of social media from Personal to Academic activities, arguably forming a basis of operation irrespective of activity. For instance, while a Baby Boomer might seek the library to review a particular field of study or alternatively create an account on a social media site to access this information, a Generation X or Y might use their existing accounts and facilities to quickly retrieve the same information with ease and convenience, forming part of his or her regular modus operandi.

Across all three generations, exclusively Academic use is rarest. Taking a more in-depth look at individual services, it can be immediately observed that the use of Facebook and Video Sharing platforms feature equally between people using them exclusively for Personal use or for both. Both for Generation X and Y, the most popular tool is seen as video sharing services with 59.63% and 61.29% claiming to utilize them for academic purposes respectively. This is followed by Facebook, Twitter, Blogs and Google+ in Generation Y, while Generation X is employing LinkedIn, followed by Facebook, Blogs, Twitter and Google+. Other services show a much lower degree of penetration here, such as Photo Sharing or Bookmark Sharing. It could be argued that these services are already covered by the social networking sites, such as Facebook, and this is reflected here.

The participants were asked if and how often they produce and share certain types of content on social media platforms in order to determine their activities. These were personal status update, re-sharing other's content, sharing links, own blog/text, own video, and own photo. Across generational groups results appear consistent, that the majority of respondents are social media consumers rather than producers, with BB on the extreme end of non-producers and Generation Y closer to the producers end as shown in Fig. 5.

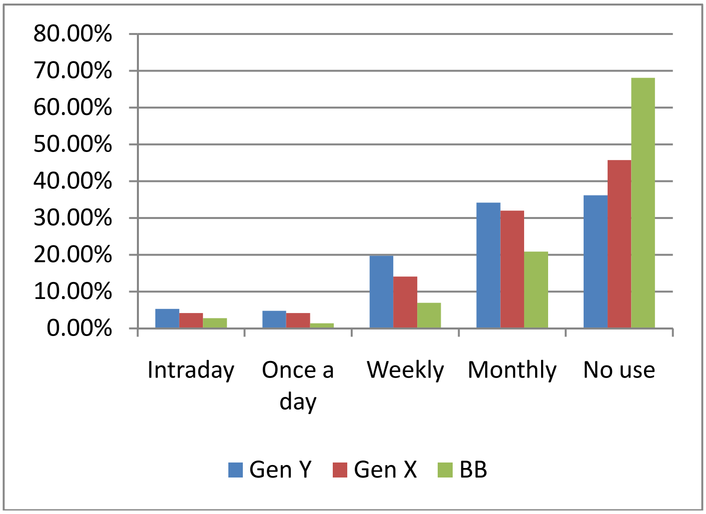



**Fig. 5.** Content creation and sharing in social media platforms

### 3.4 Opinions and Beliefs about Social Media

The following analysis focuses on the perception of social media platforms per generational group as summarized in Table 3 and Table 4. Specifically, they look at the need to be on social platforms, the ease of use of social media platform, the feeling of security they provide and the impact on social association felt by the user.

With respect to the necessity of having access to and being active on social media platforms, irrespective of generational grouping, one can observe a fair degree of consensus.

When considering the ease of use of social media platforms, overall respondents considered social media platforms easy to use and useful. However, we see that when they were asked if they find it difficult to follow the information flow on numerous social media platforms, there is a moderate trend towards a perceived difficulty as we move from the Generation Y to the BB for cross-platform use. Interestingly, the majority of respondents across generations considered it difficult (Completely Agree or Slightly Agree) to follow information flows there, which is equally reflected in the general difficulty to use multiple platforms. In summary, it appears that generational differences play no to very little part in the perception of social media platforms with respect to their usefulness, and a universal difficulty is encountered when trying to manage information across several.

**Table 3.** Perception of social media platforms (Q1: It is necessary to use SM in today's world, Q2: It is important to be on SM platforms, Q3: I find it difficult to get used to changes/new things on SM platforms, Q4: I only use SM platforms for leisure time, Q5: SM platforms are useful, Q6: I find it difficult to follow the information flow on numerous SM platforms, Q7: I find it difficult to actively use numerous SM platforms, Q8: SM platforms are easy to use.)

| **Generation Y** | **Q1** | **Q2** | **Q3** | **Q4** | **Q5** | **Q6** | **Q7** | **Q8** |
|---|---|---|---|---|---|---|---|---|
| Completely Agree | 57% | 33% | 6% | 6% | 46% | 22% | 23% | 57% |
| Slightly Agree | 35% | 42% | 21% | 26% | 50% | 37% | 39% | 39% |
| Slightly Disagree | 6% | 20% | 44% | 42% | 3% | 29% | 29% | 3% |
| Completely Disagree | 2% | 5% | 28% | 25% | 1% | 13% | 9% | 0% |
| **Generation X** | **Q1** | **Q2** | **Q3** | **Q4** | **Q5** | **Q6** | **Q7** | **Q8** |
| Completely Agree | 56% | 41% | 7% | 2% | 40% | 39% | 32% | 44% |
| Slightly Agree | 32% | 39% | 36% | 21% | 56% | 41% | 42% | 50% |
| Slightly Disagree | 11% | 14% | 44% | 49% | 3% | 14% | 19% | 6% |
| Completely Disagree | 1% | 6% | 13% | 28% | 1% | 6% | 6% | 1% |
| **Baby Boomers** | **Q1** | **Q2** | **Q3** | **Q4** | **Q5** | **Q6** | **Q7** | **Q8** |
| Completely Agree | 33% | 33% | 17% | 0% | 8% | 58% | 50% | 8% |
| Slightly Agree | 33% | 42% | 33% | 42% | 67% | 33% | 25% | 67% |
| Slightly Disagree | 25% | 17% | 42% | 33% | 17% | 8% | 25% | 25% |
| Completely Disagree | 8% | 8% | 8% | 25% | 8% | 0% | 0% | 0% |

Further, we consider the perceived risks of social media platforms across the generational groupings. Answers to a questionnaire item about controlling who can access the content added to social media platforms indicate the feeling of control respondents have over the shared information on social media decreases as we move from Generation Y to BB, with a clear decreasing trend from 80% feeling in control to just 33% respectively, with Generation X still at 63%. Interestingly however, when queried on the associated privacy concerns and security of social media platforms, all 3 generations are dubious and consider them largely insecure. Extrapolating from this, one might argue that we therefore merely observe a difference in the risk-tolerating behaviour of the generational groups, with the older generations being more risk averse. This behavioural pattern is in line with the characteristics of the BB as studied and documented by several researchers in literature[4].

[4] Generations in History, `http://www.lifecourse.com/assets/files/gens_in_history(1).pdf`, by Life Course Associates, last accessed in April 2014.

**Table 4.** Perception of social media platforms ( Q9: I can control who can access the content I add to SM platforms, Q10: SM platforms have security risks regardless of how informed the users, Q11: SM platforms are risky for personal information and privacy, Q12: I do not think that my SM accounts can be hacked by others, Q13: Being on SM platforms makes me feel like I am not alone, Q14: SM platforms make me feel like I belong to a group, Q15: I feel uncomfortable unless I check my SM accounts at least once a day, Q16: Being on SM platforms makes me feel uneasy.)

| **Generation Y** | **Q9** | **Q10** | **Q11** | **Q12** | **Q13** | **Q14** | **Q15** | **Q16** |
|---|---|---|---|---|---|---|---|---|
| Completely Agree | 41% | 43% | 25% | 11% | 10% | 3% | 19% | 7% |
| Slightly Agree | 38% | 44% | 52% | 25% | 35% | 10% | 29% | 28% |
| Slightly Disagree | 14% | 12% | 20% | 43% | 34% | 38% | 28% | 38% |
| Completely Disagree | 6% | 2% | 3% | 21% | 21% | 49% | 24% | 27% |
| **Generation X** | **Q9** | **Q10** | **Q11** | **Q12** | **Q13** | **Q14** | **Q15** | **Q16** |
| Completely Agree | 24% | 46% | 39% | 8% | 7% | 2% | 14% | 8% |
| Slightly Agree | 39% | 48% | 50% | 26% | 37% | 16% | 25% | 38% |
| Slightly Disagree | 28% | 6% | 10% | 39% | 39% | 41% | 27% | 32% |
| Completely Disagree | 10% | 1% | 0% | 28% | 17% | 41% | 35% | 22% |
| **Baby Boomers** | **Q9** | **Q10** | **Q11** | **Q12** | **Q13** | **Q14** | **Q15** | **Q16** |
| Completely Agree | 8% | 75% | 50% | 8% | 8% | 17% | 0% | 8% |
| Slightly Agree | 25% | 25% | 50% | 8% | 25% | 8% | 17% | 33% |
| Slightly Disagree | 42% | 0% | 0% | 42% | 25% | 50% | 25% | 25% |
| Completely Disagree | 25% | 0% | 0% | 42% | 42% | 25% | 58% | 33% |

Lastly, we consider the feeling of association the use of social media triggers. Broadly, we observe that irrespective of generation, though users feel not alone or isolated when engaged with social media, they, however, do not feel any association to any social grouping.

About half of BB apparently experience a sense of unease when interacting with social media, a statistic equally observable in Generation X users, with marginally fewer in Generation Y. Strangely, when queried for a felt need to check social media on a daily basis, almost 50% of respondents in Generation Y felt compelled to engage in social media with a quick drop to 39% in Generation X and just 17% for BB.

### 3.5 Leveraging Education with Social Media

Participants were asked if they have ever used any social media platform for education. Not surprisingly, social media use for educational purposes is highest amongst Generation Y (who are mostly students and technologically able young faculty members) at 87% and Generation X (who are mostly faculty members that are more open to utilizing technology for learning than their predecessors) at 71% respectively, with a dramatic drop thereafter at 33% for the BB who are introduced to the digital technologies much later in their lives.

Next we investigate the positive or negative impact that social media is perceived to have in education according to our findings. Most notably perhaps, BB and Generation X appear to be attributing general usefulness to social media in education, while Generation Y respondents predominantly disagreed with this assertion.

In surprising contrast, the responses of Generation X on the other questions were predominantly affirming of the positive impact social media can have in education and considered it not to decrease their performance either, in line with that of Generation Y and BB. Overall, there appeared general consensus on these aspects. Two differences to note are that Generation Ys consider the beneficial impact on ICT skills to be limited, with only 50% considering it beneficial compared to 85% by Generation Xs. This could be explained as the Generation Ys usually see themselves as technological and with advanced ICT skills, whereas the older generations might be feeling more insecure about their ICT skills or that there is room for improvement. Similarly, an increase in productivity is only believed by 40% in Generation X compared to an immense 82% in Generation Y. This could also be enlightened in a similar explanation, where groups who are insecure about their ICT skills think that integrating social media could be a costly process in terms of effort and time, and therefore having a negative impact on their performances.

## 4 Discussion and Conclusion

This research could indicate ways of if and how researchers, faculty members and university administrations could adjust their practices to strengthen technology enhanced learning by employing social media tools in higher education. Our study focused on the current use and perception of social media in general and for education in a developing country, where we aimed to find answers for *readiness* and *willingness* of the students and instructors regarding the idea of using social media in general and for educational purposes; their *awareness* of educational use of social media; and if and how these *differ across generations*.

Our dataset contained faculty members and students, with the majority from Generation Y, less in Generation X and very little from the Generation BB (which makes it hard to draw general conclusions for BB). At the moment Generation Z is not in higher education and therefore did not partake in our study. Therefore the resulting demographic distribution was expected for any higher education institution in the present, but one should keep in mind that Generation Z is soon to join the picture. Hence, when strategies for technology enhanced learning are to be decided, these demographics and the future shift should be considered.

We observe that although Generation X is closer to Generation Y in terms of awareness of social media and willingness to change in a linear spectrum, they still show signs of hesitations and concerns. Given the trend between generations and the fact that Generation Z are born into the Internet age, it would be safe to assume that they will lead the spectrum in social media usage, followed by Generation Y and X and extending the gap between Generation Z and BB in terms of technological demands. Although one might argue that Generation BB is soon to retire and therefore disappear from the picture, an enforced strategy change might make this group uncomfortable. Thus it is important to consider this fact when educational policies and strategies are being developed.

While almost all of Generation Ys are frequent users of social media followed by Generation X by three quarters, we see that Generation BB range between regular users and seldom users, defined as a few times a month. We can say that for Generation Y and X, where the graphics show more intense intraday and once a day use, social media is already forming a part of their daily lives.

The two most popular social media services among Generation Y and X were found to be the social networking site Facebook and video sharing platforms. Furthermore, 32% of Generation Y claimed to use Facebook for academic use while this was 28% for Generation X and only 12% for BB. For the other popular service among Generation Y and Generation X, video sharing platforms, 61% and 60% claimed to use it for academic and personal use respectively. However, we observe that when it comes to producing content for these platforms Generation Y leads by 23.19% and is followed by Generation X at 11.01%. This may indicate that older generations are more likely to be consumers in social media than the younger generations. This might lead to difficulties in educational sharing and communication on social media among generations.

Following the responses, one can conclude that social media is considered beneficial irrespective of generational association in education as well as in general. However, experiences and concerns seem to differ across generations. Although participants think that social media is easy to use and useful, BB mentioned it is difficult to manage and follow numerous social media platforms, while Generation X and Y also struggle to maintain information flow cross-platform. This finding is essential to keep in mind when a social media platform is to be utilized for education. We can say that agreeing on one or two common platforms for education can be manageable and useful for benefiting from the advantages of social media and education amalgamation, while randomly employing social media platforms in formal education might put a burden on users, which might then outweigh these benefits. Therefore, common strategies as well as institutional support seem necessary. Studying the advantages and limitations of various platforms could be the path to choosing the best common platforms, but one should keep in mind that it might be easier to adopt and integrate those platforms which are already in use by the majority.

When such generic solutions are pursued, security and privacy concerns of users should be taken into account as well. Our findings showed that older generations might need support or training for social media use in education as they mentioned that they do not feel completely in control of their social media activities or accounts.

**Acknowledgements.** This study was supported by the Scientific Research Fund of the Middle East Technical University Northern Cyprus Campus; project number SOSY13YG1. Authors would like to thank Goknur Kaplan Akilli and Tayfun Can Onuk for their assistance during the survey development and administration. Authors would also like to thank survey participants who took the time from their busy schedules to participate in the study.

## References


1. Laru, J., Näykki, P., Järvelä, S.: Supporting small-group learning using multiple Web 2.0 tools: A case study in the higher education context. The Internet and Higher Education 15(1), 29–38 (2012) ISSN 1096-7516
2. Garoufallou, E., Charitopoulou, V.: Web 2.0 in library and information science education: the Greek case. New Library World 113(3/4), 202–217 (2012)
3. Boulos, M., Maramba, I., Wheeler, S.: Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. Journal BMC Medical Education 6(1), 1–8 (2006), http://dx.doi.org/10.1186/1472-6920-6-41
4. Tess, P.: The role of social media in higher education classes (real and virtual) – A literature review. Computers in Human Behavior 29(5), A60–A68 (2013), http://dx.doi.org/10.1016/j.chb.2012.12.032, ISSN 0747-5632
5. Davis III, C.H.F., Deil-Amen, R., Rios-Aguilar, C., González Canché, M.S.: Social media and higher education:A literature review and research directions. Report printed by the University of Arizona and Claremont Graduate University (2012)
6. Roblyer, M.D., McDaniel, M., Webb, M., Herman, J., Witty, J.V.: Findings on Facebook in Higher Education: A Comparison of College Faculty and Student Uses and Perceptions of Social Networking Sites. Internet and Higher Education 13(3), 134–140 (2010)
7. Madge, C., Meek, J., Wellens, J., Hooley, T.: Facebook, social integration and informal learning at university: 'It is more for socialising and talking to friends about work than for actually doing work'. Learning, Media and Technology 34(2) (2009)
8. Johnson, K.A.: The effect of Twitter posts on students' perceptions of instructor credibility. Learning, Media and Technology 36(1), 21–38 (2011)
9. McCrindle, M., Wolfinger, E.: Generations Defined. Ethos 18(1), 8–13 (2010)
10. Keles, H.N.: A research on determining motivation profiles of Generation Y employees. Organizasyon ve Yönetim Bilimleri Dergisi 3(2) (2011) ISSN: 1309 -8039
11. Yuksekbilgili, Z.: Turkish type Y generation. Electronic Journal of Social Sciences 12(45) (2013) ISSN:1304-0278
12. Prenksy, M.: Digital Natives, Digital Immigrants. On the Horizon 9(5), 1–6 (2001)
13. Calvert, S.L., Jordan, A.B.: Children in the digital age. Applied Developmental Psychology 22, 3–5 (2001)
14. Tapscott, D.: Educating the Net Generation. Educational Leadership 56(5), 6–11 (1999)
15. Celik, I., Akilli, G.K., Onuk, T.C.: Social Media for academics: motivation killer or booster? In: Proceedings of IEEE 14th International Conference on Advanced Learning Technologies, ICALT 2014 (2014)

# A Recommender System for Students Based on Social Knowledge and Assessment Data of Competences

Oscar Chavarriaga, Beatriz Florian-Gaviria, and Oswaldo Solarte

Escuela de Ingeniería de Sistemas y Computación (EISC),
Universidad del Valle, Colombia
{oscar.chavarriaga,beatriz.florian,
oswaldo.solarte}@correounivalle.edu.co

**Abstract.** TEL Recommender systems have been used to improve experiences of students or teachers. Many such systems use information about students, such as interests, preferences, and demographic data. They also use resource metadata and ratings. The authors of this paper think that recommender systems are also valuable when implemented in online or blended courses using competence-based assessment since these systems can take advantage of social knowledge about competence development, and students' performance. By using collaborative filtering and knowledge-based techniques, it is possible to obtain recommendations from social knowledge and adapt the former to each student's performance. In this paper, the authors propose a system to recommend activities and resources that help students in achieving competence levels throughout an online or blended course. This recommender system takes into consideration experiences previously stored and ranked by former students. In order to offer successful learning advice, this recommender system analyzes the student's current competence levels against similar former students' performances. Functional test results indicate that the proposed technical approach is accurate. Moreover, these results seem to reflect that social knowledge and students' qualifications are sources of valuable recommendations for online and blended courses.

**Keywords:** Technology-enhanced learning (TEL), recommender system, collaborative filtering, social knowledge, competence development.

## 1 Introduction

Recommender systems have been included in several software applications, such as online stores, searching engines, music and video hosting web sites. Even banks granting credit make use of recommender systems [1]. Recommender systems are usually employed to take advantage of stored information in data repositories. Regarding the Technology-Enhanced Learning (TEL) area, recommender systems have been used for different purposes as follows: 1) annotating in context, 2) finding good items, 3) finding novel resources, 4) predicting and supporting students' performances, and some others [2] [3]. The most used techniques to find good

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 56–69, 2014.
© Springer International Publishing Switzerland 2014

learning items and predict student's performances are collaborative filtering and knowledge-based techniques. To provide some examples of these techniques there are novel TEL recommender systems using collaborative filtering [4] [5] as well as the ones using knowledge-based techniques [6] [7]. Furthermore, some instances of TEL recommender systems joining these two techniques are also available in [8] [9].

Recommender systems can be used as a tool for exploiting *social knowledge* stored in course-learning repositories. Adolphs in [10] describes *social knowledge* as knowledge of the minds of others. According to [1], *social knowledge* is the one about the larger community of users other than the target user. Furthermore, *social knowledge* enables the use of collaborative algorithms in which predictions about individuals are extrapolated from their community opinions [1]. Some examples of recommender systems exploiting *social knowledge* are found in [7] and [11]. The implementation of course-learning repositories such as Learning Management Systems (LMS), Massive Open Online Courses (MOOC), Training systems, and Online Informal Courses enable the collection of valuable information of former students' experiences. Some of the experience data obtained from these course-learning repositories could be pieces of advice, actions, material that students used successfully, and the students' historical performances (grades or qualifications). Hence, this information could be used as the *social knowledge* of a course-learning repository to give support to new students in this social learning environment. New students could improve their learning performance by learning from historical social experiences. However, social knowledge can grow and become overwhelming if many historical courses are taking into account. A recommender system is a tentative option to address this scenario; it can take advantage of this *social knowledge*, and filter the most relevant information for personalized recommendations.

With the previous ideas in mind, this paper will describe a TEL recommender system prototype. This particular system will use collaborative filtering techniques to assist current students in a competence-based assessment course. The suggested recommender system will take advantage of documented historical experiences and qualifications. Moreover, the system will filter the best historical experiences for any given student in relation to former students based on: (1) similar performance and (2) successful achievement of the next expected competence levels in the course. Thereby, the objectives of this system will be to find good items (former students' experience data) and support student performance improvement.

This paper is organized in eight different sections as follows. Section 2 shows related work of recommender systems on TEL. This part highlights algorithms and relevant ideas found in the studied recommender systems. Section 3 describes competence-based assessment frameworks and the technology software product where this recommender was added. Section 4 explains the recommender process. Section 5 displays the implemented prototype of this model and its testing scenario. Another ongoing testing work is also mentioned in this section. In Section 6, the preliminary findings of the authors are presented. The authors outline their conclusions in Section 7. Finally, Section 8 gives some ideas about future work.

## 2 Related Works

Most recommender systems on education are implemented based on standard recommendation approaches: content-based, collaborative-based or hybrid-based filtering. They analyze students' information and resources metadata in order to find patterns or to retrieve information that is hidden in the data. The main idea is to look for either items rated high by similar users or for items similar to those rated high by the user. There also exist cases where the data to be studied are not just students' attributes or resources attributes. In [6] a recommender system composed of subsystems is presented. Subsystems extract and process information related to psychological and physiological student profiles, taking into account aspects that can have a significant impact on learning efficiency. In the same way, the recommender system proposed in this paper uses students' assessments during a particular online or blended course to make recommendations. If psychological and physiological profiles can affect the learning efficiency and can be used to find patterns in the data in order to make recommendations, it is valid to analyze and check if the qualifications obtained in a course can be used for the same objective. These data partially represent students' performance and is representative of student in the system. In [4] authors associate IQ level with students' skills and their learning performance. Using the Stanford-Binet intelligence scale, associations between the level obtained and learning performance are searched. In a similar way, achieved competence level can be used. In [4], the system makes its decision according to the IQ level reached by a student, and retrieves recommendations. Thus, it is possible to make decisions based on the competence level achieved by students.

In some cases, educational recommender systems aim to support students at a specific time point; the presence of a time parameter affects the recommendation being retrieved. For instance, [5] proposed a recommender system to support students in selecting their undergraduate program. The quantity of missing credits required to be completed for a specific career path is one of the parameters most taken into consideration. Based on the value of this parameter, the system can change recommendations to guide students throughout the undergraduate program. This mechanism is also used in our proposal of recommender system, but on another scale. Recommendations aim to guide students throughout a specific course according to their current performance. However, recommendations may change during the course development based on the time parameter.

In educational environments, knowledge transfer is a priority, but students need to develop skills too. The European Qualifications Framework (EQF) [12] may be used for analyzing both of these. When a course is designed, it is necessary to recognize related knowledge and skill competences in order to decompose general evaluation into more specific assessments, after which, teachers can assess students both in knowledge and skills gained. Considering that a theoretical field is supported by knowledge and that a practical field is supported by skills, in [8], a recommender system to support transfer of knowledge and skills between novice worker, teacher (theory) and trainer (work area) is proposed. The system is based on the idea that information transfer in educational environments is different from work environments. The proposed recommender system uses the EQF assessments in order to help students with the achievement of knowledge competences and skill competences.

Filtering techniques are selected according to available information in data. In [9], a hybrid recommender system is proposed. The authors use Learning Objects (LO) as recommendation items; they use attributes of LO in order to find recommendations using content-based filtering. They also build a student profile with his/her attributes and preferences in order to find recommendations using collaborative filtering. The available data for the proposed recommender system in this paper are the student's qualifications. These qualifications can represent the student profile and can be analyzed using a collaborative filtering technique.

Table 1 summarizes the studied recommender systems. This table characterizes important features that are valuable, such as the type of input data, if a time parameter is used, if the recommender takes into account development of competences, etc.

**Table 1.** Summary of related works

| | Input data | | | | | Goal | | Technique | |
|---|---|---|---|---|---|---|---|---|---|
| **Recommender System** | **Using typical data (Demographic, preferences,** | **Using other type of data** | **Using time parameter that may change decisions** | **Taking into account competence development** | **Taking into account social knowledge** | **Finding good items** | **Helping to improve learning performance** | **Knowledge based** | **Collaborative Filtering** |
| [6] | | x | | | | | x | x | |
| [4] | | x | | | | | x | | x |
| [5] | x | | x | | | | x | | x |
| [8] | | x | | x | | | x | x | x |
| [9] | x | | | | | x | | x | x |

## 3 Competence-Based Learning

Worldwide, competence-based learning uses qualification models to assess student competences. Currently, some relevant qualifications frameworks in Europe are the European Qualifications Framework (EQF) [12], and the Common European Framework of Reference for Languages (CEFR) [13]. These qualification frameworks are used to design learning programs to certify student qualifications with different types of competences assessed at different levels. These frameworks are reference models of qualifications for learning programs.

All such qualification frameworks define types of competences, levels, and descriptors associated to these levels. Table 2 shows a general competence frameworks template.

**Table 2.** Competence frameworks template

| | **Competence type 1** | **…** | **Competence type n** |
|---|---|---|---|
| **Level 1** | Descriptor | Descriptor | Descriptor |
| **…** | Descriptor | Descriptor | Descriptor |
| **Level m** | Descriptor | Descriptor | Descriptor |

The EQF identifies three types of competences: Knowledge, Skills and Personal competences. This framework defines eight levels of achievement, from the basic Level 1 to the highest level 8. For each level, this framework exposes its corresponding descriptors. Every descriptor represents what a student needs to show as evidence of learning for a level and type of competence. This framework is the lifelong learning framework set for the European Commission [12].

Concerning the European Qualification Framework, it is the Common European Framework of Reference for Languages (CEFR) [13]. It is used to assess comprehension, writing, and oral expression competences of learners in any language. These are the types of competences assessed on this framework. It uses six levels to evaluate the language learning process: A1, A2, B1, B2, C1, and C2. Level A1 achieved by learners who can express and understand themselves using simple daily phrases. Level C2 achieved by learners who can easily understand everything they read or hear, and express themselves with great fluency and precision, even in complex situations.

As in Europe, the Colombian context also presents the implementation of a new competence framework. There is a new government initiative in Colombia to train basic and middle education teachers. This initiative is based on a new competence framework that assess teachers in Information and Communication Technologies (ICT) competences [14]. This framework uses six types of competences: Technological, pedagogical, instructional design, communicative, management, and research competences. The framework defines three acquisition levels: Level 1 Explorer, level 2 Integrator, and level 3 Innovator.

The recommender system proposed in this paper aims to extend [15], a software suite that helps to introduce the EQF (or any competence framework based on Table 2) into formal and informal courses. This application contributes to the learning process by supporting the design, the monitoring, the assessment, and the analytic visualization of qualifications in online or blended courses. When a course is designed, teachers can define assessment activities (AA) on it. For each AA, they define the competences that are going to be assessed and what is the expected competence level of achievement for that particular activity. When an AA is finished, teachers assess students by indicating the level they achieved on every competence relevant to the activity. Students can then be aware of what they have achieved, of what they have not achieved, or of those levels of achievement where they have exceeded expectations. At this time, students can indicate the experiences they did for achieving competence levels, and they can also score the existing experiences they followed from others (teachers or former students). This is the process in [15] that captures assessment data and recommendation items. Table 3 shows an example of expected competences levels for each AA in a course assessed using the EQF. The expression N/A appears in this table when the competence is not assessed in that activity.

The authors are interested in exploring the benefits of recommender systems in online and blended courses with competence-based assessment. The conclusion in [16] indicates that *social knowledge* about competence development can be mined by recommender systems.

**Table 3.** Expected competence levels in Linear Control Systems Course

| | Assessment activities (AA) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Competence** | $AA_1$ | $AA_2$ | $AA_3$ | $AA_4$ | $AA_5$ | $AA_6$ | $AA_7$ | $AA_8$ | $AA_9$ |
| **Ability to solve problems of linear systems** | 1 | 1 | 2 | 2 | 3 | 4 | 4 | 4 | 4 |
| **Knowledge in control of linear systems** | 1 | 1 | 2 | 2 | 3 | 4 | 4 | 4 | 4 |
| **Punctuality with assigned tasks** | 1 | N/A | 2 | N/A | N/A | 3 | N/A | 3 | N/A |

## 4 Recommender Process

This system has been designed to extend the AEEA Suite [15], a system that helps to design, monitor, evaluate, and analyze courses in higher educational environments, under an usual competence framework template (See Table 2).

In competence-based courses, a recommender system that uses assessment data in order to assist students through courses is welcome. This model uses a collaborative filtering technique in order to retrieve recommendations.

The goal of this recommender system is to assist students throughout a course, under a competence evaluation model. For each course-competence, students must achieve increasing competence levels through developmental activities. Recommendations could be useful to learners in achieving the next competence level. The main idea of the recommender system is to find similar students according to their learning curve (represented by the process of acquiring competence levels). Then, the system asks for the items they rate and scores these items according to their importance (given by the similarity between the student who gave the advice and the student who will be recommended). At the end of the process, the system retrieves the recommendations.

The system recommends that teachers and former students score different items such as pieces of advice, actions, material that students used successfully. These data are stored in the database in order to make new students' recommendations. Students can add new items in every AA. Figure 1 shows how recommendation items are scored or added to the database. Some recommendation item examples are:

- "Redo exercises on the 3rd class"
- "Watch http://www.youtube.com/watch?v=8TMBjfS8wY0"
- "Exercises 12-18 book X, p. 23"
- "Read 15-18 pages on Dynamic programming v1.pdf"

Before starting the recommendation process, the system needs to know the specific course time point; it presents recommendations to achieve incoming expected competence level going from a time point A to a time point B in a course.

Figure 2 shows the process for calculating and retrieving recommendations. This process describes how to transform former students' qualifications into recommendation items for a current student. The process is divided into five phases: Pre-processing data, filtering former student candidates, calculating similarity, ranking and retrieving recommendation items.

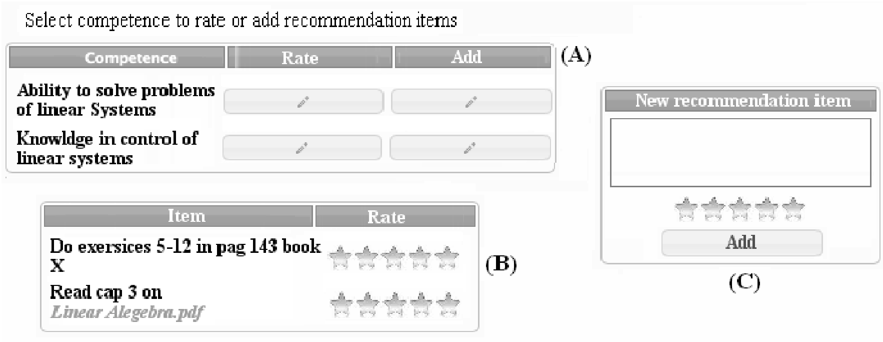



**Fig. 1.** Rating and adding recommendation items

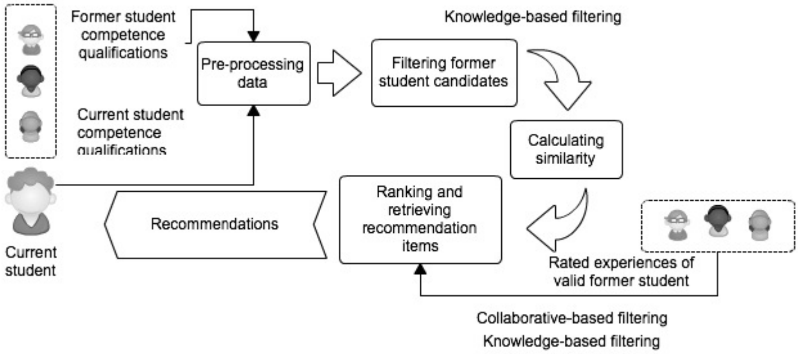



**Fig. 2.** Recommendation process

### 4.1 Pre-processing Data

The recommender system receives two input data. First, the students' qualifications linked to the course in the past. This information is represented with a matrix $m_x n$, having $m$ former students and $n$ assessment activities of the course. Assuming that the course is predefined with eight AA, Table 4 shows an example of former students' qualifications for each AA in a specific course competence. Second, up-to-date qualifications of current students are collected. After grouping these data, the recommendation process can continue to the next phase.

**Table 4.** Former students' qualifications (Historic data)

| | $AA_1$ | $AA_2$ | $AA_3$ | $AA_4$ | $AA_5$ | $AA_6$ | $AA_7$ | $AA_8$ |
|---|---|---|---|---|---|---|---|---|
| **Andy** | 1 | 2 | 2 | 2 | 3 | 3 | 4 | 4 |
| **Bob** | 1 | 2 | 2 | 3 | 3 | 3 | 3 | 4 |
| **Dylan** | 1 | 2 | 2 | 2 | 2 | 4 | 4 | 4 |
| **Alice** | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 4 |

### 4.2 Filtering Former Student Candidates

This phase filters former students who achieved the expected level in the next course AA. The idea is to remove students who did not achieve the next expected competence level.

For example, based on Table 4, and an expected level of 3 on the $AA_5$, the process to remove invalid candidates is to check students who did not achieve the expected competence level on the $AA_5$. According to Table 5, Dylan was removed from dataset.

**Table 5.** Filtering students who did achieve the goal

| | $AA_1$ | $AA_2$ | $AA_3$ | $AA_4$ | $AA_5$ | Filter Result |
|---|---|---|---|---|---|---|
| **Andy** | 1 | 2 | 2 | 2 | 3 | Approved |
| **Bob** | 1 | 2 | 2 | 3 | 3 | Approved |
| **Dylan** | 1 | 2 | 2 | 2 | **2** | **Removed** |
| **Alice** | 1 | 2 | 3 | 3 | 3 | Approved |

## 4.3 Calculating Similarity

The goal in this phase is to calculate similarity between current student and former students filtered in the previous phase. Qualification coincidences in their learning curve give the similarity. The Jaccard's coefficient [17] is used to calculate similarity between students and is given by Equation 1.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

Based on the data in Table 5, for a current student with qualifications {1, 2, 2, 2} in activities 1, 2, 3, and 4, respectively, the qualification coincidences are shown in Figure 3.

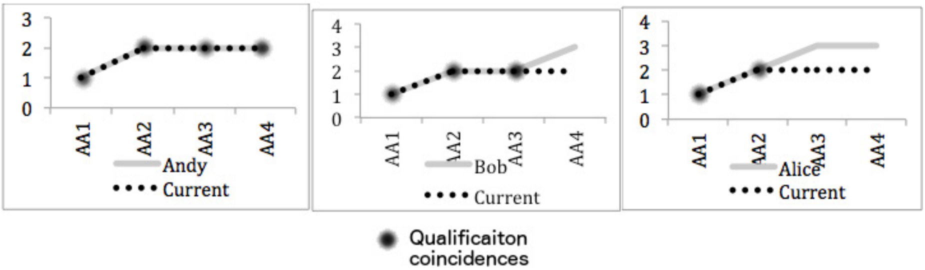



**Fig. 3.** Qualification coincidences

Equation (1) is used to calculate similarity between current student and former students. Table 6 shows the Jaccard's coefficient between current student and former students. Andy's performance has the highest similarity to current student.

**Table 6.** The Jaccard's similarity coefficient

| | $AA_1$ | $AA_2$ | $AA_3$ | $AA_4$ | Similarity |
|---|---|---|---|---|---|
| **Current** | 1 | 2 | 2 | 2 | |
| **Andy** | 1 | 2 | 2 | 2 | 1 |
| **Bob** | 1 | 2 | 2 | 3 | ¾ |
| **Alice** | 1 | 2 | 3 | 3 | ½ |

### 4.4 Ranking and Retrieving Recommendation Items

Recommendation items are previously rated and stored by former student. When a student rates some item, this action is stored in database. Table 7 shows possible ratings of four items ($I_1$, $I_2$, $I_3$, $I_4$). The goal of this phase is to automatically rank recommendation items. Ranking is based on the Jaccard's coefficient previously calculated and the ratings of items stored in the database.

**Table 7.** Previously rated recommendation items

| Student | Andy | Andy | Bob | Bob | Alice | Alice | Alice |
|---|---|---|---|---|---|---|---|
| **Item** | $I_1$ | $I_2$ | $I_1$ | $I_2$ | $I_1$ | $I_3$ | $I_4$ |
| **Rate** | 4 | 3 | 4 | 4 | 5 | 2 | 3 |

Equation 2 gives the ranking of an item $I_n$. Here, $k$ is the number of former students who rated item $I_n$, $JC_j$ is the Jaccard's Coefficient between the student $j$ and current student, and $R_j$ is the rating that student $j$ gives to item $I_n$.

$$I_n = \sum_{j=1}^{k}(JC_j * R_j) \tag{2}$$

Using Equation 2 and based on Table 6 and Table 7 data, the ranking for items $I_1$, $I_2$, $I_3$ and $I_4$ is:

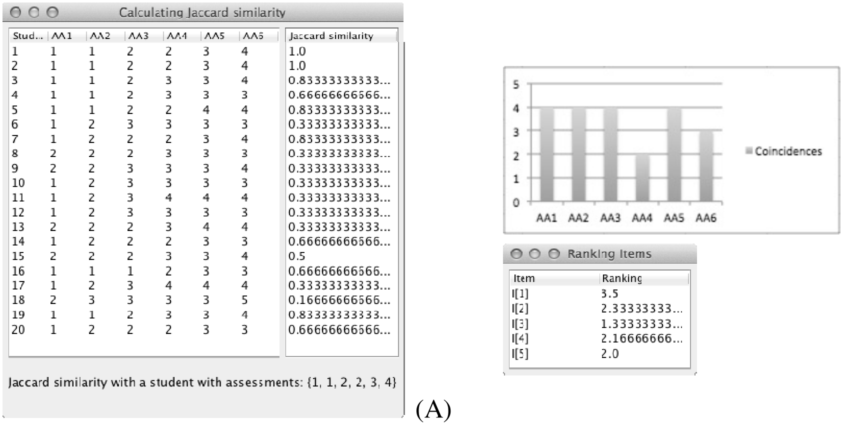
$I_1$ = (1*4)+(¾*4)+(½*5) = 7,5
$I_2$ = (1*3)+(¾*4) = 4
$I_3$ = (½*2) = 1
$I_4$ = (½*3) = 1,5

After ranking the recommendation items, the system sorts out items by relevance. In this example, the relevance order is $I_1$, $I_2$, $I_4$, and $I_3$. Finally, the most relevant items are shown for the current student.

# 5 Prototype

The recommender system prototype was implemented using the Java programming language and a PostgreSQL database. Recommendation items and qualifications were generated synthetically, with an intentional focus on results. Obtained results correspond to expected results, and the functional tests applied to recommender system were satisfactory. Figure 4, Figure 5, and Figure 6 show different applied tests to assess the correctness of the recommender's functionality. There were 20 former students and 5 recommendation items in the dataset. In the 3 scenarios, for simplicity, the value was the same for each item rating and enabled us to observe the changes in the results due to the collaborative filtering and the recommender restrictions. Students 1-4 rated the first item, students 5-8 rated the second item, students 9-12 rated the third item, students 13-16 rated the fourth item, and students 17-20 rated the fifth item.

First scenario: Figure 4(A) shows the 1st scenario, where the first group of 4 former students has the highest similarity with the current student according to their learning curve. For each AA, Figure 4(C) shows the number of coincidences in qualifications between the current student and the former students group.




| Stud... | AA1 | AA2 | AA3 | AA4 | AA5 | AA5 | Jaccard similarity |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 2 | 2 | 3 | 4 | 1.0 |
| 2 | 1 | 1 | 2 | 2 | 3 | 4 | 1.0 |
| 3 | 1 | 1 | 2 | 3 | 3 | 4 | 0.83333333333... |
| 4 | 1 | 1 | 2 | 3 | 3 | 3 | 0.66666666666... |
| 5 | 1 | 1 | 2 | 2 | 4 | 4 | 0.83333333333... |
| 6 | 1 | 2 | 3 | 3 | 3 | 3 | 0.33333333333... |
| 7 | 1 | 2 | 2 | 2 | 3 | 4 | 0.83333333333... |
| 8 | 2 | 2 | 2 | 3 | 3 | 3 | 0.33333333333... |
| 9 | 2 | 2 | 3 | 3 | 3 | 4 | 0.33333333333... |
| 10 | 1 | 2 | 3 | 3 | 3 | 3 | 0.33333333333... |
| 11 | 1 | 2 | 3 | 4 | 4 | 4 | 0.33333333333... |
| 12 | 1 | 2 | 3 | 3 | 3 | 3 | 0.33333333333... |
| 13 | 2 | 2 | 2 | 3 | 4 | 4 | 0.33333333333... |
| 14 | 1 | 2 | 2 | 2 | 3 | 3 | 0.66666666666... |
| 15 | 2 | 2 | 2 | 3 | 3 | 4 | 0.5 |
| 16 | 1 | 1 | 1 | 2 | 3 | 3 | 0.66666666666... |
| 17 | 1 | 2 | 3 | 4 | 4 | 4 | 0.33333333333... |
| 18 | 2 | 3 | 3 | 3 | 3 | 5 | 0.16666666666... |
| 19 | 1 | 1 | 2 | 3 | 3 | 4 | 0.83333333333... |
| 20 | 1 | 2 | 2 | 2 | 3 | 3 | 0.66666666666... |

Jaccard similarity with a student with assessments: {1, 1, 2, 2, 3, 4}

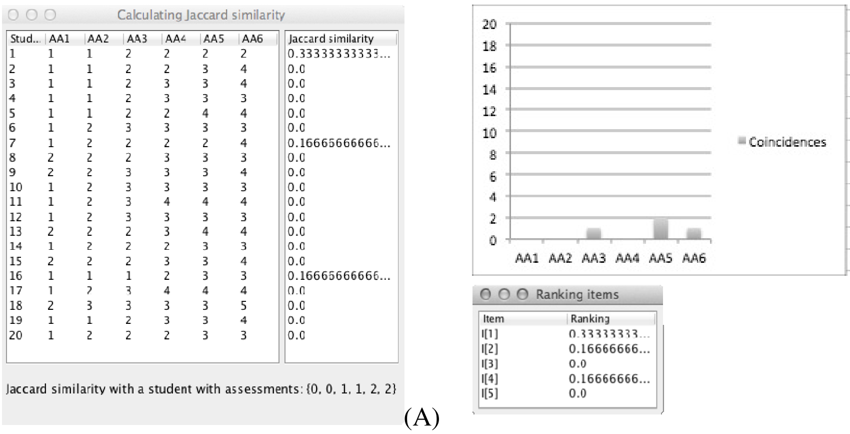
| Item | Ranking |
|---|---|
| I[1] | 3.5 |
| I[2] | 2.33333333... |
| I[3] | 1.33333333... |
| I[4] | 2.16666666... |
| I[5] | 2.0 |

(A) (B) (C)

**Fig. 4.** First test scenario

The first group of former students rated (I[1]), and as might be expected, this recommendation item had the highest ranking - see Figure 4(C); therefore, it would be the first option to recommend.

Second scenario: Figure 5(A) shows the 2nd scenario; in this scenario, the current user has a poor similarity to all of the former students' learning curves. For each AA, Figure 5(B) shows the number of coincidences in qualifications between the current student and the entire former student group.




| Stud... | AA1 | AA2 | AA3 | AA4 | AA5 | AA6 | Jaccard similarity |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0.33333333333... |
| 2 | 1 | 1 | 2 | 2 | 3 | 4 | 0.0 |
| 3 | 1 | 1 | 2 | 3 | 3 | 4 | 0.0 |
| 4 | 1 | 1 | 2 | 3 | 3 | 3 | 0.0 |
| 5 | 1 | 1 | 2 | 2 | 4 | 4 | 0.0 |
| 6 | 1 | 2 | 3 | 3 | 3 | 3 | 0.0 |
| 7 | 1 | 2 | 2 | 2 | 2 | 4 | 0.16666666666... |
| 8 | 2 | 2 | 2 | 3 | 3 | 3 | 0.0 |
| 9 | 2 | 2 | 3 | 3 | 3 | 4 | 0.0 |
| 10 | 1 | 2 | 3 | 3 | 3 | 3 | 0.0 |
| 11 | 1 | 2 | 3 | 4 | 4 | 4 | 0.0 |
| 12 | 1 | 2 | 3 | 3 | 3 | 3 | 0.0 |
| 13 | 2 | 2 | 2 | 3 | 4 | 4 | 0.0 |
| 14 | 1 | 2 | 2 | 2 | 3 | 3 | 0.0 |
| 15 | 2 | 2 | 2 | 3 | 3 | 4 | 0.0 |
| 16 | 1 | 1 | 1 | 2 | 3 | 3 | 0.16666666666... |
| 17 | 1 | 2 | 3 | 4 | 4 | 4 | 0.0 |
| 18 | 2 | 3 | 3 | 3 | 3 | 5 | 0.0 |
| 19 | 1 | 1 | 2 | 3 | 3 | 4 | 0.0 |
| 20 | 1 | 2 | 2 | 2 | 3 | 3 | 0.0 |

Jaccard similarity with a student with assessments: {0, 0, 1, 1, 2, 2}

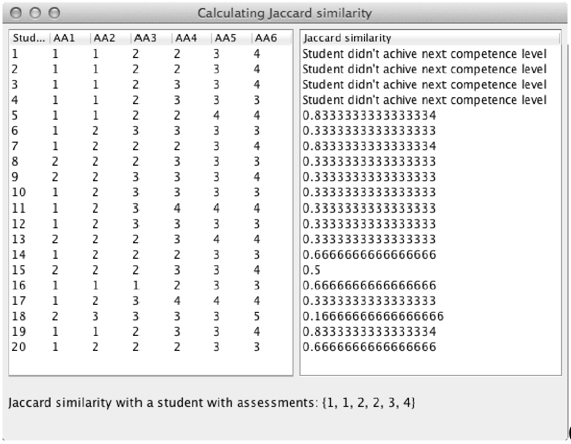
| Item | Ranking |
|---|---|
| I[1] | 0.33333333... |
| I[2] | 0.16666666... |
| I[3] | 0.0 |
| I[4] | 0.16666666... |
| I[5] | 0.0 |

(A) (B) (C)

**Fig. 5.** Second test scenario

The ranking values obtained are the lowest relative to the previous scenario - see Figure 5(C). Then, the system can decide which criterion to use, such as select the first options in the ranking, or define a threshold in order to select only the recommendation items that exceed this threshold.

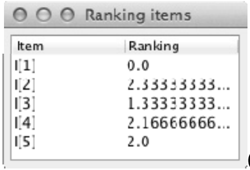
Third scenario: In this scenario, all the qualifications data were the same as that used in the 1st scenario, but the students who rated the first item did not achieve the next expected competence level. Figure 6(A) shows this scenario. Then, the first item received a lower ranking value. Thus the system changes its decisions; its first option to retrieve would now be I[2], see Figure 6(B).


Calculating Jaccard similarity

| Stud... | AA1 | AA2 | AA3 | AA4 | AA5 | AA6 | Jaccard similarity |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 2 | 2 | 3 | 4 | Student didn't achive next competence level |
| 2 | 1 | 1 | 2 | 2 | 3 | 4 | Student didn't achive next competence level |
| 3 | 1 | 1 | 2 | 3 | 3 | 4 | Student didn't achive next competence level |
| 4 | 1 | 1 | 2 | 3 | 3 | 3 | Student didn't achive next competence level |
| 5 | 1 | 1 | 2 | 2 | 4 | 4 | 0.8333333333333334 |
| 6 | 1 | 2 | 3 | 3 | 3 | 3 | 0.3333333333333333 |
| 7 | 1 | 2 | 2 | 2 | 3 | 4 | 0.8333333333333334 |
| 8 | 2 | 2 | 2 | 3 | 3 | 3 | 0.3333333333333333 |
| 9 | 2 | 2 | 3 | 3 | 3 | 4 | 0.3333333333333333 |
| 10 | 1 | 2 | 3 | 3 | 3 | 3 | 0.3333333333333333 |
| 11 | 1 | 2 | 3 | 4 | 4 | 4 | 0.3333333333333333 |
| 12 | 1 | 2 | 3 | 3 | 3 | 3 | 0.3333333333333333 |
| 13 | 2 | 2 | 2 | 3 | 4 | 4 | 0.3333333333333333 |
| 14 | 1 | 2 | 2 | 2 | 3 | 3 | 0.6666666666666666 |
| 15 | 2 | 2 | 2 | 3 | 3 | 4 | 0.5 |
| 16 | 1 | 1 | 1 | 2 | 3 | 3 | 0.6666666666666666 |
| 17 | 1 | 2 | 3 | 4 | 4 | 4 | 0.3333333333333333 |
| 18 | 2 | 3 | 3 | 3 | 3 | 5 | 0.16666666666666666 |
| 19 | 1 | 1 | 2 | 3 | 3 | 4 | 0.8333333333333334 |
| 20 | 1 | 2 | 2 | 2 | 3 | 3 | 0.6666666666666666 |

Jaccard similarity with a student with assessments: {1, 1, 2, 2, 3, 4}

(A)


Ranking items

| Item | Ranking |
|---|---|
| I[1] | 0.0 |
| I[2] | 2.33333333... |
| I[3] | 1.33333333... |
| I[4] | 2.16666666... |
| I[5] | 2.0 |

(B)

**Fig. 6.** Third test scenario

Figure 7 shows how the recommendation items are shown to the current user.

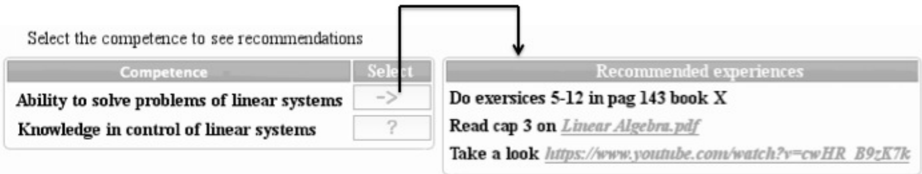



**Fig. 7.** Showing recommendation items

# 6 Findings

Information, such as demographic data, preferences, interests or behavior, is not the only data that can be the basis of a TEL recommender system. In this paper, it has been shown that it is possible to design and implement a TEL recommender system based on documented historic experiences and competence assessments of former students.

Functional tests confirmed that the recommender process proposed is suitable to assist incoming students with recommendations that former students had stored and rated.

Testing results ratify that this hybrid TEL recommender system using collaborative filtering techniques and knowledge-based techniques allow taking advantage of social knowledge in online and blended courses with competence-based assessment as a basis for recommendations.

Preliminary presentations of this recommender system to the community of TIT@ training course [18] and the CIER-SUR alliance [19] [20] [21], suggest that the University of Valle, the Secretary of Education in Valle-Colombia, and the Ministry of National Education in Colombia are interested in hosting the assessment application and the recommender system application in their online platforms to assess and train basic and middle school teachers in ICT.

Further tests are needed, but the authors think that this TEL recommender system can help students to be better informed in taking decisions at every step of their online or blended courses with competence-based assessment because they could learn from others' experiences.

## 7 Conclusions

In this paper, a recommender-system prototype for online and blended courses with competence-based assessment was presented. The system could help to improve the incoming students' performance because the former students have stored and rated several of their good experiences that helped them succeed. This system uses collaborative filtering techniques and knowledge-based techniques to make use of the experience and output of former students. Therefore, activities and resources can be suggested to help current students to obtain and improve towards expected competence levels planned for the course.

This recommender system encourages *social knowledge*, due to fact that the participants of a course can contribute ideas that can be used by other students in the future. Hence, incoming students will be more able to obtain expected competence levels (planned by teachers) in their own learning processes. The authors' approach also exploiting teachers' opinions about activities and related resources as their documented outlooks are also taking into account in the assessment application.

This recommender system extends the AEEA Suite [16] [15] functionality. The assessment tool was improved by adding methods to collect former students' successful experiences and ratings. When students either achieved or exceeded expected competence levels in an assessment, they were encouraged to store their experiences for future learners.

The approach proposed in this document can be used by the TEL community in online or blended courses for different contexts of learning (informal and formal learning). Because many competence qualifications frameworks follow the same pattern of definitions (i.e., descriptors in a matrix of type of competences vs. competence levels), this approach may be used to implement new recommenders for courses using some of those competence frameworks.

Preliminary functional testing, using the limit value technique and synthetic data, validated the technical soundness of the recommender process. It is necessary to use the recommender in real courses, to collect a large source of social knowledge and overcome the early adoption problem of knowledge-based recommender systems. To do so, the software suite has been used in a series of massive blended courses for training teachers of basic and middle public schools in Cali, Colombia as is described in the next section.

## 8 Future Work

This is an ongoing research work. To validate whether this recommender system improves students' performance in competence development, a series of further testing is planned. By the time this paper is completed, the software suite (AEEA Suite) is been used in blended massive courses for the TIT@ training plan [18]. In TIT@, at least 500 local teachers of basic and middle public schools are been trained on ICT competences. These teachers are divided into successive groups taking the same course. In this ongoing work, the suite was configured to use the framework [14] instead of the EQF. Then, the recommender system of this paper is going to be used in the training course of the CIER-SUR alliance [19] [20] [21]. The latter course will train other 3000 teachers of basic and middle schools of the Colombian South Region on ICT competences. Similar to the TIT@ course, in the CIER-SUR course, teachers will be divided into groups taking the same course successively.

Regarding improvements for this recommender system, the authors have some ideas. Firstly, classification of former students' experiences would help students to filter recommendations by choosing categories (good resources experiences, good activities experiences, good online material, etc.) in the displayed data. Secondly, the social learning analytics tool of the software suite can be extended with visual analytics of the recommendations. A final planned extension is a recommender system for teachers regarding course design and students in danger of either dropping out or failing the course.

**Acknowledgements.** The authors would like to thank Universidad del Valle, Colombia for funding the research project NUBE-UV (CI 2756). This work has been sponsored as part of the NUBE-UV project. In addition, the authors specially thank Professor Martha Millán for her valuable consulting on the Knowledge Discovery in Databases area.

## References

1. Ricci, F., Rokach, L., Bracha, S., Kantor, P.B. (eds.): Recommender Systems Handbook. Springer, US (2011)
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, pp. 387–415. Springer, US (2011)

3. Manouselis, N., Drachsler, H., Verbert, K., Duval, E.: Recommender Systems for Learning, New york, Heidelberg, Dordrecht, London (2012)
4. Arockiam, L., Charles, S., Arul Kumar, V., Cijo, P.: A Recommender System for Rural and Urban Learners. In: Nagamalai, D., Renault, E., Dhanuskodi, M. (eds.) CCSEIT 2011. CCIS, vol. 204, pp. 619–627. Springer, Heidelberg (2011)
5. Cho, J., Kang, E.: Personalized Curriculum Recommender System Based on Hybrid Filtering. In: Luo, X., Spaniol, M., Wang, L., Li, Q., Nejdl, W., Zhang, W. (eds.) ICWL 2010. LNCS, vol. 6483, pp. 62–71. Springer, Heidelberg (2010)
6. Kaklauskas, A., et al.: Recommender System to Analyse Students' Learning Productivity. In: Yang, D. (ed.) Informatics in Control, Automation and Robotics, vol. 2. LNEE, vol. 133, pp. 161–164. Springer, Heidelberg (2011)
7. Carrer-Neto, W., Hernández-Alcaraz, M.L., Valencia-García, R., García-Sánchez, F.: Social knowledge-based recommender system. Application to the movies domain. Expert Syst. Appl. 39, 10990–11000 (2012)
8. Kong, X., Boll, S., Heuten, W.: A Hybrid Multi-recommender System for a Teaching and Learning Community for the Dual System of Vocational Education and Training. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 613–614. Springer, Heidelberg (2013)
9. Andrea, P., Marín, R., Morales, V.T.: Modelo multi-agente para recomendación híbrida de objetos de aprendizaje Multi-agent Model for Hybrid Recommendation of Learning Objects Modèle multi- agent pour recommandation hybride d ' objets d ' apprentissage Mots-clés. Rev. Virtual Univ. Católica Del Norte 40, 96–111 (2013)
10. Adolphs, R.: The social brain: neural basis of social knowledge. Annu. Rev. Psychol. 60, 693–716 (2009)
11. Wan, X., Jamaliding, Q., Okamoto, T.: Analyzing Learners' Relationship to Improve the Quality of Recommender System for Group Learning Support. J. Comput. 6, 254–262 (2011)
12. European Commission: The European Qualifications Framework for Lifelong Learning (EQF) (2008)
13. Heyworth, F.: The Common European Framework. ELT J. 60, 181–183 (2006)
14. Calderón, G.G., Buitrago, B., Alcira Acevedo, M., Tobón, M.I.: Competencias TIC Para el Desarrollo Profesional Docente
15. Florian-Gaviria, B., Glahn, C., Gesa, R.F.: A Software Suite for Efficient Use of the European Qualifications Framework in Online and Blended Courses. IEEE Trans. Learn. Technol. 6, 283–296 (2013)
16. Florian-Gaviria, B.: Technology-Enhanced Supoort For Lifelong Competence Development in Higher Education (2013), http://www.tdx.cat/handle/10803/101517
17. Maimon, O., Rokach, L.: Data Mining and Knowledge Discovery Handbook (2010)
18. TIT@ Educación Digital para Todos, https://www.facebook.com/educaciondigitaltita
19. Centros de Innovación Educativa, Construyendo Capacidades en Uso de TIC para Innovar en Educación, http://www.colombiaaprende.edu.co/html/productos/1685/w3-propertyvalue-48190.html
20. Un CIER para la investigación Las Universidades del consorcio CIER-SUR presentan sus proyectos de investigación, http://dintev.univalle.edu.co/saberabierto/adjuntos/saberabierto_ed30.pdf
21. Centro de Innovación Educativa Regional Sur, http://cms.univalle.edu.co/ciersur/

# An Evaluation Framework for Data Competitions in TEL

Hendrik Drachsler[1], Slavi Stoyanov[1], Mathieu d'Aquin[2], Eelco Herder[3], Marieke Guy[4], and Stefan Dietze[3]

[1] Open Universiteit Nederland, Welten Institute, The Netherlands
{hendrik.drachsler,slavi.stoyanov}@ou.nl
[2] The Open University, KMI, United Kingdom
Mathieu.Daquin@open.ac.uk
[3] L3S, Research Center, LUH Hannover, Germany
{herder,dietze}@l3s.de
[4] Open Knowledge Foundation, OKF, United Kingdom
marieke.guy@okfn.org

**Abstract.** This paper presents a study describing the development of an Evaluation Framework (EF) for data competitions in TEL. The study applies the Group Concept Method (GCM) to empirically depict criteria and their indicators for evaluating software applications in TEL. A statistical analysis including multidimensional scaling and hierarchical clustering on the GCM data identified the following six evaluation criteria: 1.Educational Innovation, 2.Usability, 3.Data, 4.Performance, 5.Privacy, and 6.Audience. Each of them was operationalized through a set of indicators. The resulting Evaluation Framework (EF) incorporating these criteria was applied to the first data competition of the LinkedUp project. The EF was consequently improved using the results from reviewers' interviews, which were analysed qualitatively and quantitatively. The outcome of these efforts is a comprehensive EF that can be used for TEL data competitions and for the evaluation of TEL tools in general.

**Keywords:** Evaluation Framework, Assessment of TEL tools, Data competition, Group Concept Mapping.

## 1 Introduction

With the raise of data science, there is also a new wave of publications on learning analytics, personalisation and adaptation techniques. But these data-driven research approaches in education are hardly comparable with each other. Most of the reported experiments in the TEL world are not comprehensible and do not provide the underlying research data, neither they describe their learner model, educational reasoning or personalisation techniques in sufficient detail to repeat an experiment. This is a mayor challenge for TEL science, as it is the nature of science to gain general knowledge that can be reproduced under controlled conditions. It is of crucial importance to overcome this issue to make TEL research results more reliable for the community as well as for policy makers and funding bodies. We should strive towards a comprehensive knowledge base about the effects of TEL tools on learning and teaching. For this

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 70–83, 2014.
© Springer International Publishing Switzerland 2014

reason, the dataTEL Theme Team funded by the STELLAR Network of Excellence identified the need for a comprehensive Evaluation Framework as one of the Grand Challenges for technology enhanced learning [1]. In TEL, evaluation has specific characteristics as it needs to take into account technical and educational measures. The technical measures guarantee that the software is working properly and the educational measures indicate the impact of the technology on learning scenarios.

The paper presents an empirical study for the development of an Evaluation Framework for the LinkedUp project that organises three consecutive data competitions in TEL - Veni, Vidi and Vici [2]. The goal of the LinkedUp Veni competition was to gather innovative and robust tools that analyse and/or integrate large scale, open Web data for educational purposes. Veni was open for any end-user application that analyses and makes use of Linked Data or Open Web Data for online learning. The EF has been applied to compare data-driven tools in TEL and rank them according to their achievements in a standardised manner.

The paper will report about this process by first reviewing related evaluation approaches in TEL. Then we introduce the Group Concept Mapping method that was used by the TEL community to identify evaluation criteria and indicators that are suitable for TEL related data competitions. Third, we provide some demographic data about the participants of the study and describe the procedure in further details. Fourth, we present the results of the Group Concept Mapping and summarize its main findings. Fifth, we discuss the first version of the LinkedUp EF as it has been applied to the LinkedUp competition– Veni[1]. Finally, we report how the EF was evaluated and further improved for the Vidi competition.

## 2 Related Work

### 2.1 Data Competitions

Data competitions are of increasing importance as a mean to gain knowledge about data science in various domains. Data competitions enable the data owners to review diverse approaches towards a single dataset and are therefore a strong instrument for innovation purposes.

There are various competitions related to the TEL domain such as the *Elsevier Grand Challenge*[2], where the goal is to improve communication of scientific information. The *Semantic Web Service Challenge*[3] is another example of similar initiative that aims at evaluating Semantic Web Services Mediation, Choreography and Discovery technologies. Collaborative Annotation of a Large Biomedical Corpus (*CALBC*[4]) is a European Support Action addressing the automatic generation of a large, community-wide shared text corpus annotated with biomedical entities. Evaluation will be

[1] http://www.linkedup-challenge.eu

[2] http://www.elseviergrandchallenge.com/

[3] http://sws-challenge.org

[4] http://www.calbc.eu/

performed against the harmonized contributions that have been gathered from the participants' contributions to the same challenge.

Although previous examples have demonstrated the suitability and usefulness of organised challenges and competitions to drive innovation, there are still some issues that limit the reusability and impact. Most of these initiatives are *technology-centric rather than outcome-centric*, are based on *artificial, limited test data* and *often lack of real-world scenarios.* One serious drawback in the previous competitions was lack of information about evaluation frameworks used, particularly how the criteria and indicators were identified and with what kind of methods.

In LinkedUp we aimed to overcome those issues by creating TEL related data competitions following a *realistic use-case scenarios,* involving a *large-scale testbed of Web datasets*, and a *transparent evaluation framework* to ensure a high-level of innovation and reusability of project results in education.

## 2.2 Evaluation Approaches in TEL

A systematic literature review that we conducted identified a reasonable number of studies on evaluation of TEL tools and e-learning courses but very little information was returned on evaluation frameworks, criteria and indicators for assessing educational software applications in competitions. Most of the research on TEL evaluation has been focused on usability. Some studies combine usability with specific performance measures for learning of end users (learners). For instance, [3] modified the Nielsen's protocol for the evaluation of an e-learning program. [4] presented a comprehensive usability study that brings together end-user assessments and expert inspections, thus providing a detailed students', teachers' and experts' feedback. [5] developed an integrated model with six dimensions: learners, instructors, courses, technology, design, and environment to evaluate the satisfaction from using an eLearning tool. In another study [6] proposes an usability framework that integrates web usability and instructional design parameters and proposes motivation to learn as a new type of usability dimension in designing and evaluating e-learning applications.

Next to those usability studies there have been different approaches for the evaluation of personalisation and adaption of TEL tools that are also relevant for our EF as they require data to provide their adaptation services. [7] suggested an approach to decompose the adaptation process into two layers that are evaluated separately. This is needed because a 'monolithic' evaluation cannot provide sufficient information at a level of granularity that can be valuable for the system designer to decide which part of the system needs improvement. The layered evaluation approach is still a summative evaluation with two phases rather a formative evaluation process. Simultaneously, two other modular evaluation frameworks have been proposed. The process-based framework presented by [8] consisted of four evaluation layers, the second framework has been presented by [9] and is more detailed in terms of different components involved in the adaptation process. It also addressed the question about methods and tools appropriate for the evaluation of different adaptation modules to yield input for the development process. A merged version of the two frameworks was finally proposed and has been explored by several studies that evaluated adaptive systems [10].

Another evaluation approach has been suggested by the RecSysTEL community [11]. They propose an evaluation method for Recommender Systems in TEL by using reference datasets to make the findings of the data studies more comparable to each other. They proposed a set of references datasets that could be used to gain comparable evaluation results [13]. Several studies followed this approach since it was mentioned [12] and started to contribute evidences for comparable evaluation results.

Summarising the insights from the related work section we can concluded, that there are various approaches to evaluate TEL tools and that usability is a very common criterion. It also appeared that among usability there is a lack of transparency of used evaluation criteria and how they are operationalized in suitable indicators.

One objective of the LinkedUp project is to address this lack of transparency and develop a framework that can be applied for various domains within the TEL field. The evaluation framework we are aiming for has at least three important differences to the studies discussed. First, it is *not focused on the end user or system designer*. It rather needs to support a *jury of judges to come up with an accurate, comprehensive and transparent assessment* about a submitted tool. Second, it needs to check *if a TEL tool is technically sound* but also *innovative from an educational perspective*. Third, the evaluation cannot run over a longer time period, in fact the jury needs to be able to *make a decision about the submitted tools in a timeframe of 1-2 hours*.

## 3 Method

We applied the Group Concept Mapping (GCM) method to address the lack of a transparent EF with community-driven quality indicators within TEL [14]. The aim of the GCM is to develop an evaluation framework that is driven by high profile experts from the whole TEL community, rather than a proposal of a single research group. GCM is a structured, mixed approach applying both quantitative and qualitative measures to objectively identify an expert group's common understanding about a particular issue, in our case the evaluation indicators for open educational data. The method involved the participants in three activities: 1. Idea generation, 2. Sorting of ideas into groups, and 3. Rating the ideas on some values (Priority and Applicability of the indicators). The participants work individually, but it is the advanced statistical techniques of multidimensional scaling and hierarchical cluster analysis that quantitatively aggregate individual input of the participants to reveal shared patterns in the data. One of the distinguishing characteristics of GCM is the visualisation, which is a substantial part of the analysis. Visualisation allows for grasping at once the emerging data structures and their interrelationship to support decision-making.

### 3.1 Participants

In total, 122 external experts have been identified for the GCM study. The candidates were selected according to two criteria: (a) holding a PhD degree and (b) a publication list that demonstrates experience in developing and evaluating data- driven applications in TEL. 74 experts responded positively to the invitation to participate in the study. They registered to the GCM tool for online data collection by creating a username and password. All participants gave their research informed consent. Of all participants assigned to the study, 57 contributed to the idea generation phase, 26 completed the sorting and 26 finished the rating. Figure 1 shows an overview of the participation of the experts, who agreed to participate. A meta-analytical research including 69 GCM study suggests that 20-30 participants is the optimal number for sorting the ideas [14].

| | assigned | started | finished | checked |
|---|---|---|---|---|
| Project | 73 | 73 | *na* | *na* |
| Questions | 2 | | | |
| Brainstorming | 57 | 57 | | |
| Sorting | 44 | 42 | 26 | 26 |
| Priority | 44 | 31 | 26 | 26 |
| Applicability | 44 | 29 | 26 | 26 |

**Fig. 1.** Response rate of external experts to the LinkedUp GCM study

### 3.2 Procedure

As mentioned earlier the procedure consisted of three phases, namely: 1. Idea generation, 2. Sorting of ideas into groups, and 3. Rating on two values (priority and applicability). Afterwards the researchers analyse the data and interpret the results. The results from the GCM were then used for determining evaluation indicators, criteria and potential methods to measure the indicators for the EF.

All participants were fully informed about the purpose, the procedure, and the time needed for completing the activities. The participants were provided with a link to the brainstorming page of a web-based tool for data collection and analysis. They were asked to generate ideas completing the following trigger statement:

> *"One specific indicator of the evaluation framework for assessing the Open Web Data application in the educational domain is ..."*

During the idea generation phase, the 57 experts contributed a total of 212 original ideas. After cleaning these statements from analogical and vague ideas, and splitting the statements that contained more than one idea we were left with a list of 108 indicators. The final list of 108 indicators was randomised and sent back to the participants. In the next step they were asked to first sort the ideas into groups based on their similarity, giving a representative name to the group, and, second, to rate them on two values – *priority* and *applicability* for the use in the EF.

## 4 Results

### 4.1 Point Map of the 108 Quality Indicators

Figure 2 shows the first outcome of the multidimensional scaling analysis – a point map. The closer the statements to each other, the closer in meaning they are, which also means that more participants cluster them together. Multidimensional scaling assigns each statement a bridging value, which is between 0 and 1. The lower bridging value means that a statement has been grouped together with statements around it; e.g. statements 6, 19, 77, 89, 100, 105 on the right side of figure 2. A higher bridging value means that the statement has been grouped together with some statements further apart from either side (e.g. statement 21 or 86 in the centre of the point map). Some groups of ideas can be detected by eye inspection, but to make the process more efficient a hierarchical cluster analysis was applied.

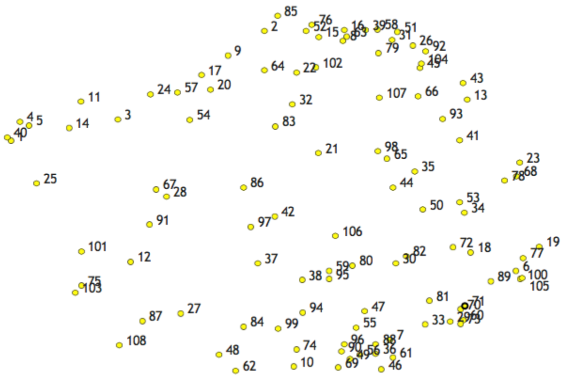


**Fig. 2.** Point map of the 108 quality indicators contributed by the TEL community

## 4.2 From the Point Map to a Cluster Map

Several solutions suggested by the hierarchical cluster analysis have been trialed (see Figure 3). For the final decision, we adapted the practical heuristic of '15-to-4' [15] as the average number of clusters per participant was 10. We started from a 15-cluster solution with the idea to arrive at a 4-cluster solution.

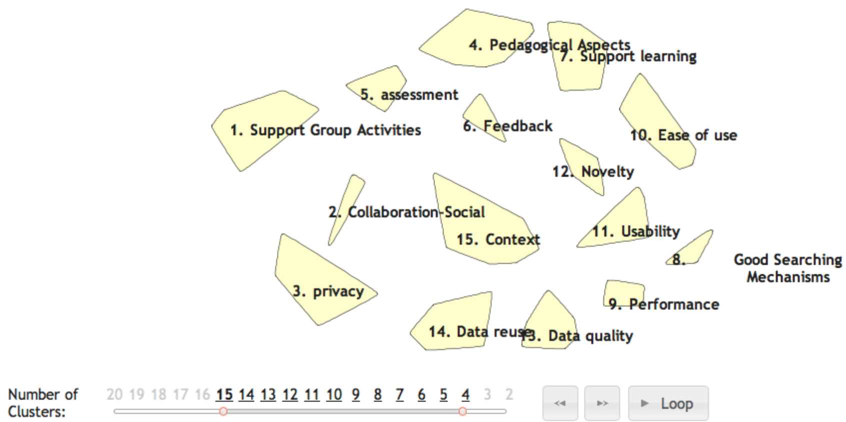



**Fig. 3.** A replay scaling 15-to-4 cluster solutions, currently shown 15 clusters

At each step, we checked whether the merging of clusters made sense for the purpose of the LinkedUp project. The six-cluster solution seemed best representing the data and serving the purpose of the study (see Figure 4).

From Figure 4 it can be seen that there is a very stable *Data* (south on the map) and *Education* (north) cluster in the point map that do not share any statements. By contrast, *Performance*, which also includes some Human Computer Interaction statements, is naturally positioned between the *Data* and *Education* clusters. The *Privacy* (west) cluster always remained apart from the other clusters, but it is also a very stable and therefore important entity for the evaluation criteria. Surprisingly, the *Support Group Activities* cluster never merged with the *Educational* clusters, as the external experts see these statements semantically different to the educational aspects of the evaluation criteria. Moreover, it developed as an additional application domain, next to the educational one, which promotes its own indicators for Open Web Data applications.

The next step of processing the clustering results is constructing meaningful labels for the clusters, using the three available methods. The first one is to check what the GCM system suggests. The second way is to look at the bridging values of the statements composing a cluster. The statements with lower bridging value represent better a cluster. The third method is to read through all statements in a cluster and define what is the story behind it. To define the clusters (criteria) we combined the three methods. We finally, chose the following labels for the 6-cluster solution: 1. Support Group Activities, 2. Privacy, 3. Educational Innovation, 4. Usability, 5. Performance, and 6. Data (see Figure 4).

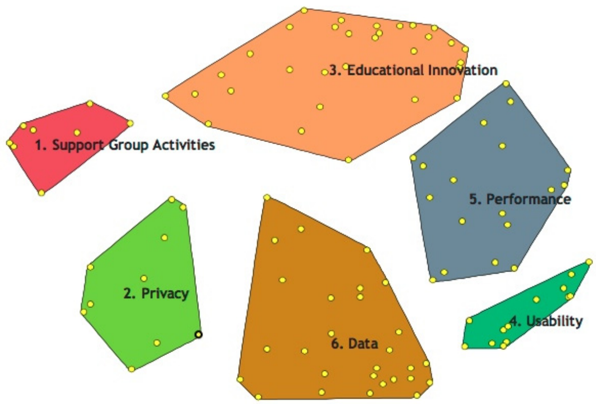



**Fig. 4.** Cluster labels

### 4.3 Six Cluster Rating Maps

As described above, the experts applied a rating to the evaluation criteria and their indicators according to two aspects of the LinkedUp EF: *Priority* and *Applicability*. *Priority* refers to the importance of a particular cluster for the evaluation of TEL tools. *Applicability* indicates the perceived ease to apply the indicator and criterion in the review process. Five layers indicate a high rating within the GCM tool, one layer of a cluster visualizes a low rating.

As Figure 5 shows, the clusters 'Usability' received the highest rating on priority followed by 'Educational Innovation' and 'Data' with three layers each. 'Support Group Activities' and 'Privacy' received the lowest score (one and two layers respectively).

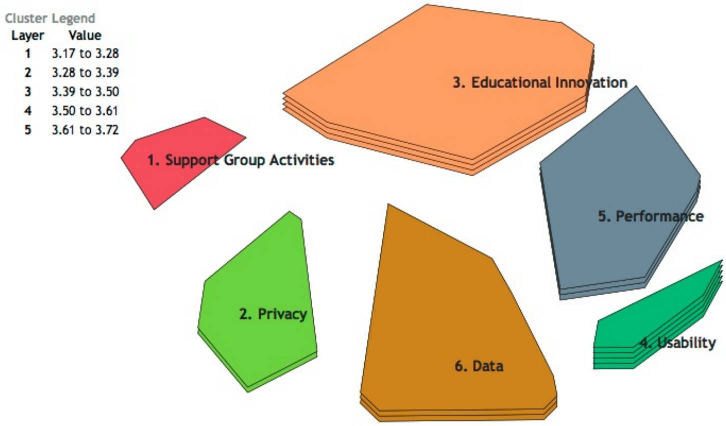



**Fig. 5.** Rating map on *priority* of the evaluation criteria/indicators for the EF.

A different picture appears for the *Applicability* aspect of the evaluation criteria (see Figure 6). According to the participants, the indicators that are easiest to implement are within the cluster '*Support Group Activities*' and *'Usability'* (four layers). *'Performance'* and *'Privacy'* are both rated with three layers as reasonably applicable indicators of the EF.

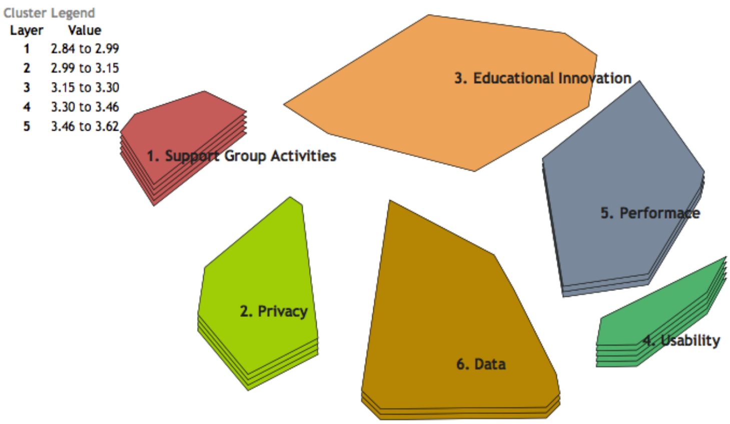



**Fig. 6.** Rating map on applicability of the evaluation criteria/indicators for the EF

The *'Educational Innovation'* cluster, which has received the highest score on priority, got very low rating here, meaning this is expected to be the most difficult to assess by the judges of a data challenge.

The ladder graph in Figure 7, called pattern match, compares the clusters on the *Priority* and *Applicability* ratings. The lines show how pairs of clusters are related according to their rating values. A Pearson product-moment correlation coefficient (r = -0.16) indicates a weak negative relationship between the two values: priority and applicability. The cluster 'Support group activities' has the biggest margin between the two values. It scores the lowest on priority and the highest on applicability. In contrast, 'Educational innovation' scores relatively high on priority but the lowest on applicability. 'Usability' scores high on both values. There is a relatively small difference between priority and applicability in the clusters 'Data' and 'Performance'.

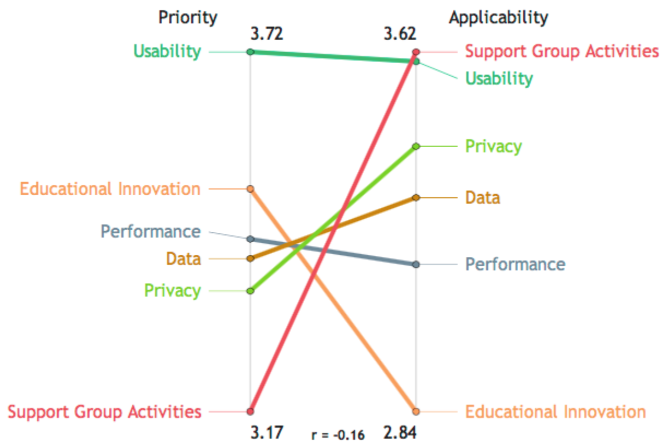



**Fig. 7.** Pattern match Priority vs. Applicability of the evaluation criteria for the EF

## 5 GCM outcomes - A First Outline of the LinkedUp EF

A panel of LinkedUp project experts discussed the results of the study in the context of the competitions that are going to be organized and decided on the final set of clusters (criteria), which included: Educational Innovation, Privacy, Usability, Performance, and Data. The sixth cluster, 'Support Group Activities' was disregarded for the first version of the EF, because it mainly contains a list of very specific features that are related to the computer-supported collaborative working (CSCW) field. It could be used later on for the specific Focus Track within a data competition around a specific CSCW use case, but was not relevant to the objectives of the first Open Call of the LinkedUp challenge.

While analysing the bridging values of all statements and their clusters, the consortium also discovered an important omission which is highly relevant for the objectives of public data competitions. The LinkedUp project aims to promote applications with high impact that constitute powerful examples of how to use Linked Data to serve different stakeholders in education. This means that applications that aim for a very narrow target group are less relevant than applications targeting a broader audience. There are some highly rated statements that can be combined in a cluster 'Audience'. Representative statements in support of such a cluster are: "That it addresses a broad community of users", or "it can be used or tailored to a variety of target groups", and "the calculation of basic metrics on technology usage (like amount of users, browsing sessions, avg. sessions per user)". The project consortium, therefore, decided to add 'Audience' as an additional criterion for the initial version of the LinkedUp EF.

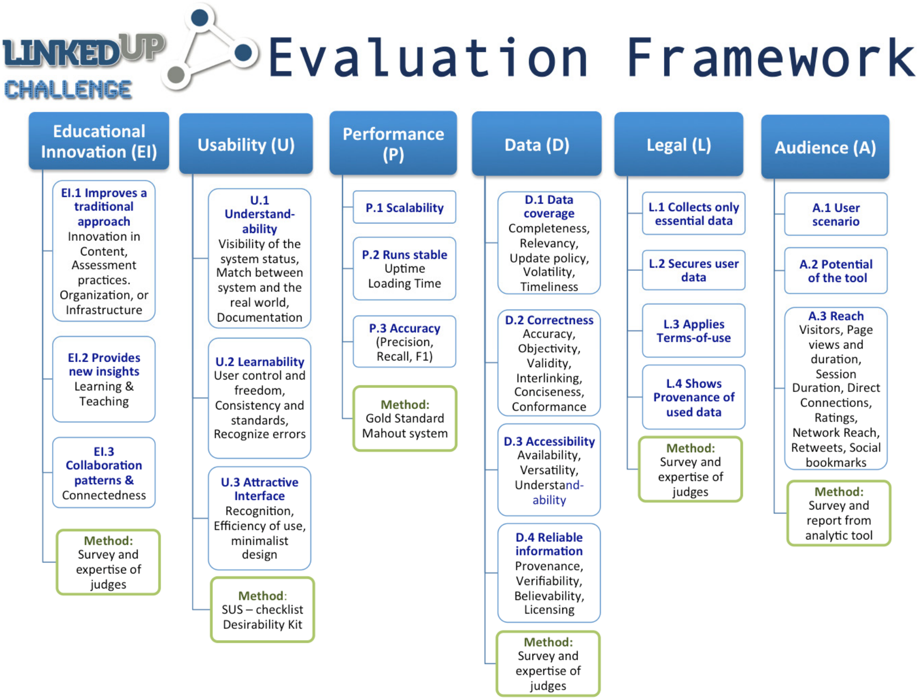



**Fig. 8.** A first version of the LinkedUp EF after the GCM study

Figure 8 represents the first version of the EF that was applied to the LinkedUp Veni competition in 2013. It shows the six criteria clusters identified by the GCM study, a set of indicators for each criteria and possible methods to measure those indicators. The indicators have been contributed by the experts participating in the brainstorming of the GCM study. The methods have been identified by another literature review for measuring the evaluation indicators that have been suggested.

## 6 Implementation of the Evaluation Framework to the Veni Competition

The line chart presented in Figure 9 shows an overview of 15 from the 22 participants submitted to the Veni competition and how they got valued according to the evaluation criteria. It indicates also how submissions scored on an individual criterion. The final scores coming form the evaluation framework were used as a basis for a deliberation process conducted by LinkedUp partners. The LinkedUp team was very satisfied with the effortless ranking of the submitted tools according to the scores. It enabled the team to shortlist the submitted tools and identify the three winners of the competition without much more efforts. In the same way it also made the evaluation results transparent to the reviewers and the participants. We could provide the participants with detailed scores about their performance on each of the evaluation criteria and contrast those with the average scores of the Veni competition.

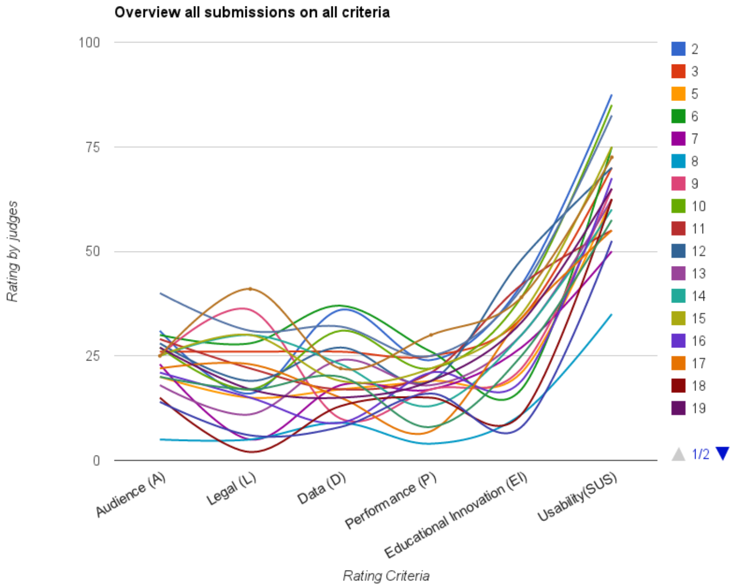



**Fig. 9.** Overview of the rating results given to the participants of the LinkedUp Veni competition. The high value of usability is affected by the chosen SUS method that has a range for the final score between 20 to 100 points.

### 6.1 Improvement of the Evaluation Framework

We conducted five individual semi-structured interviews with jury members to identify possible issues that might appear when using the EF. The jury members have been randomly selected. The interviews followed an elaborated script to provide an unified approach among the interviewers. The script contained a suggested sequence of activities, the main questions with possible probes (follow up) and how to ask questions. A letter of inform consent was also part of the document. The interview and the deliberation discussion were transcribed in verbatim before the analysis. The data analysis included both qualitative and quantitative methods.

In general the reviewers were very positive about the evaluation framework. They expressed some concerns regarding the 'not applicable' option in some of the evaluation criteria and suggested either removing this option or providing a clearer guideline how to interpret it and how to proceed with it. In addition, the reviewers felt that some of the indicators, such as 'Assessment' in the cluster 'Educational Innovation' were not relevant for the Veni competition. The reviewers suggested to have a more elaborated instruction to a criterion and possibly providing an example.

To complement the qualitative analysis of the interview data and identify issues not easily detectable in the interviews regarding the EF, we applied a free for use web text mining service called 'Text is Beautiful'[5]. The tool presents the text analytics results in the form of concept web (similar to concept map – see Figure 10). The input was the combined text of the five interviews we conducted.

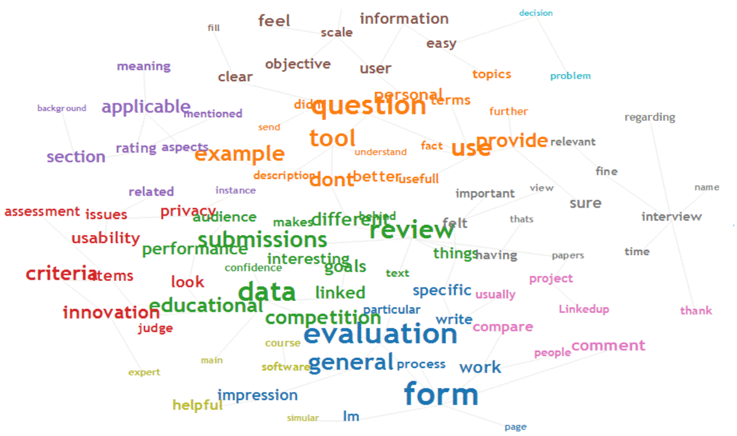



**Fig. 10.** Concept web

In the concept web, the concepts are colored-clustered in a two-dimensional space and the concepts that are strongly related are positioned in the map closely to each other. For example, the concepts in orange (e.g. 'question', 'use', 'example', 'provide', 'description', 'better', 'useful', 'description', 'understand') confirm the finding that it would be useful if an example is provided in the description of a criterion to make the instruction better formulated and understandable. The 'violet' concepts ('applicable', 'section', 'rating aspects', 'meaning', 'mentioned') reflect the discussion on the meaning of the rating aspect "not applicable". It seems to be one of the issues (see the concepts in red), specifically related to some of the criteria such as educational innovation and usability. Apparently, assessment as one of the indicators of the criteria educational innovation is an issue ('assessment issue'). The conclusion that can be drawn is that the quantitative text mining analysis had confirmed the findings from the qualitative text analysis regarding the issues that need to be addressed for the next round of the competition (e.g. the issues with the non-applicable option, issues related to the indicator 'assessment', and the need for providing an example to the instruction of each item).

[5] `http://textisbeautiful.net/`

## 7 Conclusion and Further Work

In this article we presented the findings of a Group Concept Mapping study for empirically identifying the set of criteria and indicators for evaluating data competitions in TEL. We tested the first version of the LinkedUp EF during the Veni Competition[6]. The Veni competition required '*an innovative and robust prototype or demo that used linked and/or open data for educational purposes*'. By the closing date, 22 valid submissions had been received from 12 different countries. The LinkedUp judges rated the submitted tools according to concrete criteria and indicators of the EF. The EF enabled the LinkedUp team to shortlist the submitted tools and identify the three winners of the Veni competition applying an unified and empirically validated evaluation framework. We believe that it would have been much more difficult to agree on these results without the application of the EF.

On the basis of this initial version of the EF, we are further investigating suitable evaluation criteria and their specific indicators. We are especially interested in additional metrics, and weighting to evaluate the defined criteria and automated or, semi-automated evaluation tools that can easily be applied by the LinkedUp judges saving their time.

The outcomes of the EF study could be beneficial for organisers of data competitions, not only in TEL. The EF is flexible, specific criteria and their indicators can be selected and combined to address the needs of future data competitions. We will further evaluate and improve the EF during the LinkedUp Vidi and Vici competitions to provide the most comprehensive EF by the end of the LinkedUp runtime. We are aiming to provide a kind of toolbox that can guide data scientists in setting up data competitions and hackathons in their specific domains. The toolbox will provide valuable information, lessons-learned from the LinkedUp competitions, links to suitable public data sources, hints on legal issue and how to solve those, marketing strategies, and also the EF with specific guides and templates' how to organise an accurate and transparent evaluation process for the judges and participants of data competitions. A first exploitation of this work will be done in the context of the EATEL SIG dataTEL, and the Dutch SURF SIG on Learning Analytics by organising specific data competitions as satellite projects of LinkedUp.

**Acknowledgements.** This work was partly funded by the LinkedUp (GA No:317620) and DURAARK (GA No:600908) projects under the FP7 programme of the European Commission.

## References

1. Drachsler, H., Verbert, K., Manouselis, N., Vuorikari, R., Wolpers, M., Lindstaedt, S.: Preface (Special issue on dataTEL - Data Supported Research in Technology-Enhanced Learning). Int. J. Technol. Enhanc. Learn. 4(1/2) (2012)

[6] `http://linkedup-challenge.org/veni.html`

2. d'Aquin, M., Dietze, S., Herder, E., Drachsler, H.: Using linked data in learning analytics. eLearning Papers. Nr. 37 (2014), http://www.openeducationeuropa.eu/en/elearning_papers, ISSN: 1887-1542
3. Benson, L., Elliot, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, S., Reeves, T.: Usability and Instructional Design Heuristics for E-Learning Evaluation. In: Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Denver, Colorado, USA, pp. 1615–1621 (2002)
4. Granic, A., Cukusic, M.: Usability Testing and Expert Inspections Complemented by Educational Evaluation: A Case Study of an e-Learning Platform. Educational Technology & Societ 14(2), 107–123 (2011)
5. Sun, P.-C., Tsai, R.J., Finger, G., Chen, Y.-Y., Yeh, D.: What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. Computers and Education 50(4), 1183–1202 (2008)
6. Zaharias, P.: Usability in the Context of e-Learning: A Framework Augmenting 'Traditional' Usability Constructs with Instructional Design and Motivation to Learn. International Journal of Technology and Human Interaction 5(4), 38–61 (2009)
7. Brusilovsky, P., Eklund, J.: A study of user-model based link annotation in educational hypermedia. Journal of Universal Computer Science 4(4), 429–448 (1998)
8. Weibelzahl, S.: Evaluation of adaptive systems. In: Bauer, M., Gmytrasiewicz, P.J., Vassileva, J. (eds.) UM 2001. LNCS (LNAI), vol. 2109, pp. 292–294. Springer, Heidelberg (2001)
9. Paramythis, A., Totter, A., Stephanidis, C.: A modular approach to the evaluation of adaptive user interfaces. In: Weibelzahl, S., Chin, D.N., Weber, G. (eds.) Empirical Evaluation of Adaptive Systems. In: Proceedings of Workshop at the Eighth, International Conference on User Modeling, UM 2001, Freiburg, p. 924 (2001)
10. Brusilovsky, P., Karagiannidis, C., Sampson, D.: Layered Evaluation of AdaptiveLearning Systems. International Journal of Continuing Engineering Education and Lifelong Learning, Special issue on Adaptivity in Web and Mobile Learning Services 14(4/5), 402–421 (2004)
11. Manouselis, N., Drachsler, H., Verbert, K., Duval, E.: Recommender Systems for Learning, pp. 1–61. Springer, Heidelberg (2012)
12. Verbert, K., Manouselis, N., Drachsler, H., Duval, E.: Dataset-driven Research to Support Learning and Knowledge Analytics. International Journal of Educational Technology & Society (2012)
13. Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., Duval, E.: Dataset-driven Research for Improving Recommender Systems for Learning. In: 1st International Conference Learning Analytics & Knowledge, Banff, Alberta, Canada, February 27-March 1 (2011)
14. Rosas, S.R., Kane, M.: Quality and rigor of the concept mapping methodology: a pooled study analysis. Evaluation and Program Planning 35, 36–245 (2012)
15. Kane, M., Trochim, W.: Concept mapping for planning and evaluation. Sage Publishing, Thousand Oaks (2007)

# Which Recommender System Can Best Fit Social Learning Platforms?

Soude Fazeli[1], Babak Loni[2], Hendrik Drachsler[1], and Peter Sloep[1]

[1] Open Universiteit Nederland, Centre for Learning Sciences and Technologies, Netherlands
{Soude.fazeli,hendrik.drachsler,peter.sloep}@ou.nl
[2] Multimedia Computing Group, Delft University of Technology, Netherlands
b.loni@tudelft.nl

**Abstract.** This study aims to develop a recommender system for social learning platforms that combine traditional learning management systems with commercial social networks like Facebook. We therefore take into account social interactions of users to make recommendations on learning resources. We propose to make use of graph-walking methods for improving performance of the well-known baseline algorithms. We evaluate the proposed graph-based approach in terms of their F1 score, which is an effective combination of precision and recall as two fundamental metrics used in recommender systems area. The results show that the graph-based approach can help to improve performance of the baseline recommenders; particularly for rather sparse educational datasets used in this study.

**Keywords:** Recommender system, graph, teacher, social learning platform, social network, sparsity, performance.

## 1 Introduction

Vassileva [1] introduces social recommender systems as a practical solution to help users in finding suitable resources that can support their learning process. Social recommenders are mainly based on two methods, either Collaborative Filtering (CF) algorithms or content-based algorithms. CF algorithms pass recommendations on to a user based on the opinions of many other users and their feedback on items. They first find like-minded users and create a network of so-called nearest neighbors; then they predict an item's rating for a target user on the basis of the ratings given by the nearest neighbors to this item. Content-based algorithms are based on preferences of a user summarized in a user model. They recommend an item to a user by comparing the representation of the item's content with the user's model. Due to this content dependence, CF algorithms have been applied more widely in social recommender systems because they are more flexible and require user opinions and feedback only instead of the actual content description, as do content-based algorithms.

In this research, we focus on interactions and collaboration between users in a social learning platform developed by the eContentPlus Open Discovery Space (ODS)

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 84–97, 2014.
© Springer International Publishing Switzerland 2014

project[1]. Social learning platforms combine traditional learning management systems (LMS) with commercial social networks like Facebook to provide easy content creation, access, sharing, bookmarking, etc. Beside forums and chat communities often provided in an LMS, they let users establish more connections and improve their networks of peers. To recommend the most suitable resources to the ODS users, we decided to use collaborative filtering algorithms since they focus on the similarities and overlaps of users' social activities.

Most of the CF algorithms employ similarity measures to build the nearest neighbors network that allows a recommender algorithm to learn. Such algorithms try to find like-minded users and introduce them as nearest neighbors of a target user for whom recommendations are generated. Although this kind of CF algorithms has proven quite successful in both research and practical use cases, they rely on the full user-item matrix data. That matrix, however, is not always available, particularly not in the educational domain as it involves fewer users and fewer transactions compared to e-commerce applications [2]. This problem originates from the sparse ratings of neighbors (the sparsity problem). When rating data are sparse, users are likely to receive irrelevant recommendations. Therefore, we aim to take advantage of a graph-based approach, which extends and improves the baseline nearest neighbour CF by invoking graph search algorithms. There exist quite a few approaches on improving performance of recommenders by using graph-walking algorithms [3]–[7]. Almost all use data that include either actual social (or trust) relations between users, content features of items, or tags assigned to the items. None of these, however, were available for this study. Therefore, our main research question is:

*RQ: How to generate more accurate and thus more relevant recommendations for the users in social learning platforms by employing graph-walking methods?*

Our overall aim is to find out which recommender algorithm best fits a social learning platform such as the ODS platform. In doing so, we follow the research method proposed by Manouselis et al. for evaluating TEL recommender systems [8]. After developing the conceptual model based on a literature review and after carrying out an interview study [9], we conducted an offline data study to investigate if and how the use of a graph-walking recommenders can help to improve prediction accuracy of the recommendations made based on data collected from platforms similar to the future ODS dataset. In this paper, we present results of this offline data study. As a further step, we intend to run a user study to measure user satisfaction on the recommendations generated.

The rest of this paper is organized as follows: Section 2 gives a brief review of related studies. The proposed graph-based approach used to collect the recommendations is described in Section 3. Sections 4 and 5 present the experimental study and results, respectively. Then, we explain the practical implications of the experiment in Section 6 and conclude by giving an overview of future work in Section 7.

[1] `http://opendiscoveryspace.eu`

## 2 Related Works

Manouselis et al. [10] investigated which collaborative filtering algorithm supports multi-attribute ratings of the users within an online community composed of teachers from all over the Europe. The authors reported the results on different variations of their proposed multi-attribute collaborative filtering algorithm. Cechinel et al. [11] applied several memory-based collaborative filtering algorithms on the MERLOT repository to investigate which of the algorithms used performs best. Their study focused on evaluating the collaborative algorithms in terms of the automated quality assessment of learning resources within the MERLOT dataset. Koukourikos et al. [12] proposed to use sentiment analysis to enhance collaborative filtering algorithms. Such techniques take into account the opinions of a user on the quality of the resources before recommending resources to other users. They studied performance of their proposed approach on the MERLOT repository. Although all studies mentioned presented useful insights in applying recommender system algorithms to educational datasets, none of them dealt adequately with the sparsity problem. Verbert et al. [2] presented a dataset-driven study by testing different classical collaborative filtering algorithms on a set of educational datasets, including the Travel well, MACE, and Mendeley datasets from the dataTEL project [13]. They proposed using the implicit data of users such as tags, downloads, etc. However, their approach fails in cases in which not even the implicit data of users are sufficient to find similarity patterns between users. Manouselis et al. [14] compared the results of evaluating multi-criteria algorithms on a real dataset of original data collected from the Organic.Edunet portal and a synthetic dataset including real data plus some simulated data. They simulated how the users would have rated the learning resources and then added these simulated ratings to the real dataset. But it remains unclear in how far actual user ratings match the simulated ones.

Finally, an as yet rarely used approach in the educational domain is the state-of-the-art Matrix Factorization (MF) method. It was able to tackle the sparsity problem in ACM recommender systems research by decomposing the sparse user-item ratings matrix into two matrices using latent features of the items. Manouselis et al. [8], reported only one study on MF [15]. Thai-Nghe et al. showed that MF has the potential to take into account temporal effects such as the increasing knowledge of learners. However, they did not focus on making recommendation on learning resources.

In an attempt to improve prediction accuracy of recommendations and thus to overcome sparsity, graph-based algorithms have emerged. However, very few educational applications are known. Anjorin et al. [7] aimed to make use of the tags assigned by users in a social platform called CROKODIL. They extended an existing approach, based on the PageRank algorithm. Their study used a dataset that is similar to ODS. As with ODS, extra information about learning resources or tagging data was lacking. Therefore, they could not make use of tags and keywords assigned to the learning resources.

# 3 A Graph-Based Approach

We propose to employ graph-walking algorithms in order to improve the prediction accuracy of recommendations. Such an approach first forms a graph in which nodes are users and edges are similarity relations between users. Then, it collects recommendations for a target user by walking through the target user's neighbors.

## 3.1 Creating the Graph

We take into account a Social Index (S-index) for each user, which is inspired by the H-index and calculated using the algorithm 1. The H-index is an indicator of publications of an author. It combines information on the number of publications of some author with the number of citations [20]. Similarly, the S-index of a user *u* shows not only how many times user *u* has been selected as a neighbor, but also how much the user *u* contributed to interactions on items in common with her neighbors [6]. In this study, the S-index ranges from 1 to 100, the *similarityScore* between two users from 0 to 1. The S-index helps us to extend and improve finding like-minded users (neighborhoods). We use it for sorting list of raters of an item. This list helps us to discover new neighbors for a user, who have been excluded when walking through the created graphs but still can be a useful source of information.

**Algorithm 1** Computing S-index for user u

```
upon event (COMPUTE S-INDEX| u, NeighborsList)
    SortedNeighborsLis ← SortDescendingBySimilarityScore(NeighborsList);
    FinalNeighborsLis ← Normalize(SortedNeighborsList,MaximumSindex);
    Sinde. ← 0;
    for ( similarityScore(u,n); n in FinalNeighborsList) do
        if Sindex <= similairtyScore then
            Sindex= Sindex+1;
        else
            Break;
        end if
end for
updateSindex(Sindex);
end event
```

## 3.2 Collecting Recommendations

The graph-based approach uses a modified BFS (Breadth First Search) graph search algorithm to traverse the implicit social network created using S-index and items' raters lists. We chose BFS among the well-known walking algorithms like depth first search to first poll the direct neighbors when collecting recommendations in the created user graph. The inferred neighborhoods, therefore, are not limited to the *k* nearest neighbors only; instead we provide dynamic neighborhoods beyond *k* for each target user depending on the new neighbors the graph-based approach helps us to infer. We formalize this procedure as follows:

```
G(V,E) = CreateSocialGraph();  // V contains users
                               // E contains similarity relations between users
for all u ∈V do
    ComputeSindex(u, N);       // N contains users who have user u as their neighbor
    G(V,E  ← BFS(u, G(V,E));      // E ⊂ E' where E' contains:
                               // 1. explicit similarity relations (u,n) ∈ E and
                               // 2. new inferred relations (u, n')
    TopItem ← CollectRecommendations(u, G(V, E'));
    UpdateSindex(u,N');        // N' contains new neighbors found
    UpdateSocialGraph();
end for
```

Moreover, we followed a discounting mechanism when collecting recommendations from the neighbors who appear in the BFS result. For this, we propagate the similarity scores between users who have no direct connection yet by multiplying the interconnecting users' similarity scores. This guaranties that the inferred similarity score is always smaller than the actual values of the interconnecting edges.

## 4 Experimental Study

In order to address the research question described in Section 1, we conducted an offline experiment to compare a graph-based approach with baseline algorithms. But first, we briefly describe the classification categories of CF algorithms according to their *type* and *technique* [16].

*Type* refers to model-based and memory-based algorithms. Model-based algorithms rely on probabilistic approaches to create a model of users' preferences. Examples of model-based algorithms are neural networks, Bayesian networks, and algebraic approaches such as those using eigenvectors. Memory-based algorithms use statistical and mathematical approaches based on the users' data stored in memory. Examples are the Pearson correlation coefficient, Tanimoto-Jaccard coefficient, and Euclidean distance. In general, model-based algorithms are faster than memory-based algorithms because they develop models of users' preferences offline. However, they require a full set of users' preferences to create a user model. Moreover, model-based algorithms often prove to be costly in terms of required resources and maintenance efforts. Therefore, choosing what type of CF to use is a trade-off that depends on the use case's limitations. In this study, we use both memory-based and model-based algorithms to find out which one can best help to tackle the sparsity problem.

The *technique* of CF algorithms often refers to user-based and item-based algorithms [16]. User-based algorithms try to find patterns of similarity between users in order to make recommendations; item-based algorithms follow the same process but are based on similarity between items. Here, we are interested in both user-based and item-based CFs because we focus on users' interactions and activities.

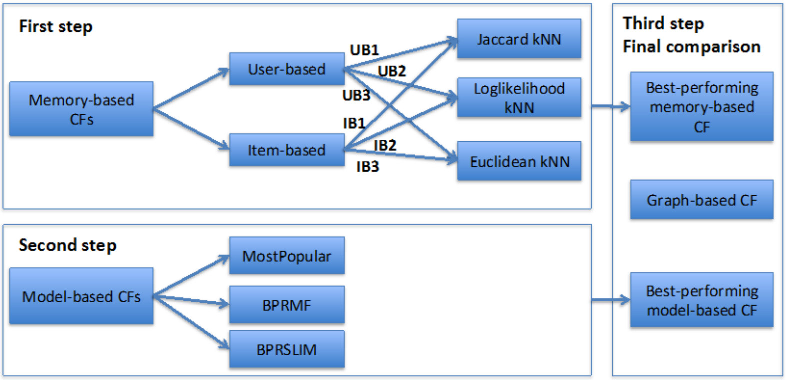



**Fig. 1.** Experimental method used in this paper

As shown in Fig. 1, the experimental study in this paper consists of three steps:

1. Memory-based CFs: we compare several both user-based and item-based k Nearest Neighbors (kNN) that can process binary data in terms of their F1 score. We choose the best-performing baseline kNN as the underlying layer for the graph-based approach.
2. Model-based CFs: we compare several model-based CFs, which are appropriate candidates for the binary data we have in this study.
3. Ultimately, we compare the graph-based approach with the best performing CFs selected from each of the previous two steps. We discuss the results to find out which of the CFs used performs best for our data and, thus, can be selected to be integrated in the ODS social platform.

In the experiment, we used Apache Mahout[2] and MyMediaLite[3] as open source frameworks; they provide implementations of the baseline collaborative filtering algorithms. Moreover, both frameworks enable us to evaluate the performance of baseline algorithms. The proposed graph-based approach has been implemented in Java.

### 4.1 Candidate Memory-Based CFs

As indicated, memory-based CF algorithms try to find similarity between users' opinions, interests, and actions on the items (from here on, we use items for learning resources). To measure similarity, we need to select a similarity measure that is able to process educational datasets including user interactions data. The educational

[2] http://mahout.apache.org/
[3] http://mymedialite.net/

datasets used in this study consist of too few explicit user preferences. But they do provide implicit user preferences such as views, downloads, tags, etc., which show users' interest in particular items as binary indicators (interested in the item: yes/no). Some of the similarity measures such as Pearson correlation are not suitable for this kind of data because they require explicit user preferences. Among the popular similarity measures, the Tanimoto-Jaccard coefficient (Jaccard coefficient, from here on), Loglikelihood, and Euclidean distance are most appropriate if the data includes implicit user preferences in binary format.

### 4.2 Candidate Model-Based CFs

Among the model-based CFs that make use of implicit user preferences, the Bayesian Personalized Ranking (BPR) method proposed by Rendel et al. [17] in our opinion best suits the data used in this study. Rendel et al. [17] aimed to optimize the learning process for the task of personalized ranking on a set of items. They applied their BPR to the state-of-the-art matrix factorization models to improve the learning process in the Bayesian model used (BPRMF) [18]. In addition, we use an extended version of BPR with Sparse Linear Methods (SLIM) [19] for item ranking optimized for BPR optimization criterion. The SLIM method [19] generates top recommendations by aggregating positive user feedback on items. This approach 'learns' a sparse aggregation coefficient matrix from aggregated users' feedback and can produce fast and accurate recommendations. Beside these two methods based on BPR, we also use the baseline MostPopular approach, which makes recommendations based on general popularity of items. In this method, items are weighted based on how often they have been seen in the past.

**Table 1.** Details of the selected datasets

| Dataset | Number of users | Number of items | Transactions | Sparsity (%) | Source |
|---|---|---|---|---|---|
| MACE | 631 | 12,571 | 23,032 | 99.7096 | MACE portal[4] |
| OpenScout | 331 | 1,568 | 2,560 | 99.5067 | OpenScout portal[5] |
| MovieLens | 941 | 1,512 | 96,719 | 93.6953 | GroupLens research[6] |

### 4.3 Datasets

We selected the MACE and OpenScout datasets for the following reasons:

- The datasets contain social data of users such as ratings, tags, reviews, etc. on learning resources. So, their structure, content and target users are quite similar to the ODS dataset.

[4] http://mace-project.eu/
[5] http://learn.openscout.net/
[6] http://grouplens.org/datasets/movielens/

- Running recommender algorithms on these datasets enables us to conduct an offline experiment for studying the recommender algorithms before going online with the actual users of the ODS.
- Both MACE and OpenScout datasets comply with the CAM (Context Automated Metadata) format [21], which provides a standard metadata specification for collecting and storing social data. CAM will also be applied in the ODS project for storing the social data.

Since the educational domain lacks a 'golden' standard dataset to run data studies, such as the MovieLens dataset, we also tested the MovieLens dataset as a reference dataset. Table 1 provides an overview of these datasets. Note that the educational datasets MACE and OpenScout clearly suffer from extreme sparsity.

# 5 Results

The offline experiment in this study gauges the F1 score. We chose F1 due to the type of users data in this study and also in typical social learning platforms, which are implicit user preferences in binary format. We could have made use of other common metrics, such as MAE, RMSE and nDCG. However, these only work if we have explicit user preferences available, like 5-star ratings This is hardly ever the case in educational settings. Another advantage of F1 is that it combines precision and recall, which are both important metrics to evaluate accuracy and coverage of recommendations generated [22]. F1 ranges from 0 to 1. In this experiment, we split users' ratings randomly into two sets: a training set (80%) and test set (20%). The sets include actual and predicted relevance indictors of users. We computed F1 for the top 10 recommendations of the result set for each user. These settings are commonly used for empirical studies on recommender systems [22].

## 5.1 Memory-Based CFs

First, we evaluated several baseline k-nearest neighbor (kNN) CFs on the similarity measures: Jaccard coefficient, Loglikelihood ratio, and Euclidean distance. We did so to find out which of them performs best on the data used in this study. Figure 1 shows the result of the F1 for testing the following baseline CFs:

1. User-based Jaccard kNN (UB1)
2. User-based Loglikelihood kNN (UB2)
3. User-based Euclidean kNN (UB3)
4. Item-based Jaccard kNN (IB1)
5. Item-based Loglikelihood kNN (IB2)
6. Item-based Euclidean kNN (IB3)

The used datasets are: MACE, OpenScout, Travel well and MovieLens. The horizontal axis (x) indicates different sizes of neighborhood ($k$) and the vertical axis (y) shows the values of F1. As Fig. 1 shows, the F1 value of the used algorithms provides different patterns depending on the used datasets. In general, user-based Jaccard kNN (UB1) provides the best F1 scores for all the datasets used: MACE (exceeding 7.7%), OpenScout (exceeding 10%), and MovieLens (exceeding 20%); only with an exception for OpenScout when k=20. Among the datasets used in Fig. 2, MovieLens owns the highest F1 values of all. The main reason for this (expected) result is the larger size of the dataset in terms of number of user transactions. The F1 result for MovieLens is consistent with the previous study by Verbert et al. [2] for user-based Jaccard kNN (UB1) (± 0.2). For both MACE and OpenScout, although UB1 has the best F1 results for all sizes of neighbors, its F1 values fluctuate while *k* increases and thus, it does not follow a clear pattern according to the sizes of neighborhoods. Unlike MACE and MovieLens, the F1 results for OpenScout quite declines by increasing *k*.

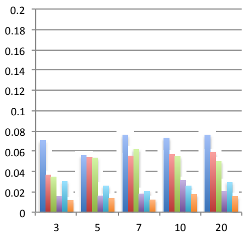



**Fig. 2.** F1 of memory-based CFs on MACE

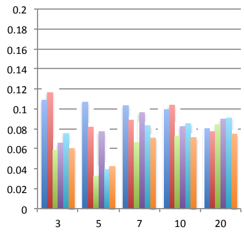



**Fig. 3.** F1 of memory-based CFs on OpenScout

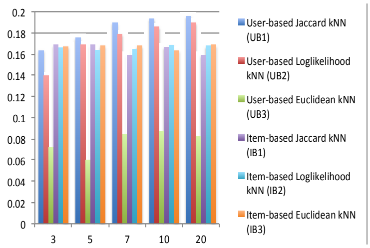



**Fig. 4.** F1 of memory-based CFs on MovieLens

F1 results of the item-based CFs are dataset-dependent; similar to the user-based CFs. Fig. 2 shows that for MACE, the user-based CFs (UB1, UB2, UB3, and UB4) outperform the item-based ones (IB1, IB2, IB3) for all *k,* with quite a large difference. The smallest difference (0.4%) is between UB2 and IB2 for *k*=3, the largest one (7.5%) is between UB4 and IB3 for *k*=20. This is unexpected since MACE has many more items (5,696) than users (105).

For OpenScout, IB1 and IB2 perform best, right after the user-based CF (UB1) for all sizes of neighborhoods and even they have better F1 than UB1's F1 when *k*=20. For MovieLens, the item-based CFs perform quite well and quite close to the best performing CFs: UB1 for the smaller sizes of neighbors (*k*=3,5). However, for *k* larger than 5, the F1 of the item-based CFs decreases compared to UB1's F1. For *k*=20, the difference between the F1 of IB3 as the best item-based CF and the best-performing user-based UB1 is more than 3%. In summery, CF recommenders that make use of similarities between users perform better than those that make use of similarities between items. We decided to select the Jaccard kNN for the ultimate comparison with the other candidate CFs.

### 5.2 Model-Based CFs

We now report the F1 results using model-based CFs on the same datasets as were used in previous sections: MACE, OpenScout, and MovieLens. As explained in section 3.1, we choose to use three model-based CFs:

1. The BPR method using Matrix Factorization (BPRMF)
2. The BPR method using SLIM (BPRSLIM)
3. MostPopular (a well-known baseline CF often used in recommender research)

Fig. 2 shows the F1 results. The horizontal axis (x) again shows the datasets and the vertical axis (y) indicates the values of F1 for the model-based methods. Similar to the results for memory-based CFs, the F1 scores of model-based CFs are also dataset-dependent and MovieLens again scores best (exceeding 11%). This refers to the lowest sparsity of this dataset compared to others (see Table1). In general, BPRSLIM performs best for all the datasets: MACE (6.76%), OpenScout (8.53%), and MovieLens (18%). BPRMF stands in the second place for MACE and MovieLens, providing F1= 5.6% and 17.7%, respectively. For OpenScout, the differences between F1 of BPRSLIM and the others are quite large (the lowest gap is around 6%). We decided to choose BPRMF since it best performs among the model-based methods used and for all the datasets.

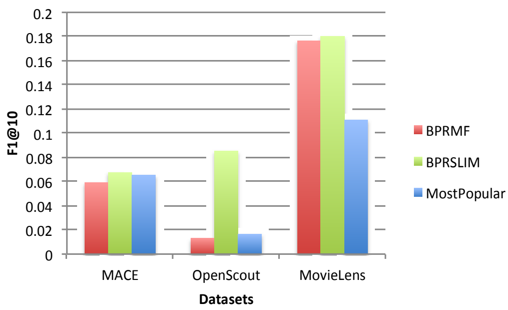



**Fig. 5.** F1 of the model-based CFs for all datasets used

### 5.3 Final Results and Discussion

In this step, we compare the graph-based approach with the best performing CFs from the previous steps: Jaccard kNN (memory-based) and BPRMF (model-based). Recall that the graph-based approach is a memory-based and user-based CF. For making a fair comparison, the graph-based CF also employs the same similarity measure used in the best performing baseline CF: the Jaccard coefficient. For the MACE and MovieLens datasets, we choose neighborhoods of size *k*=20, for which the F1 score is the highest (see Figure 1). Unlike this, for OpenScout, the best F1 score is for *k*=3.

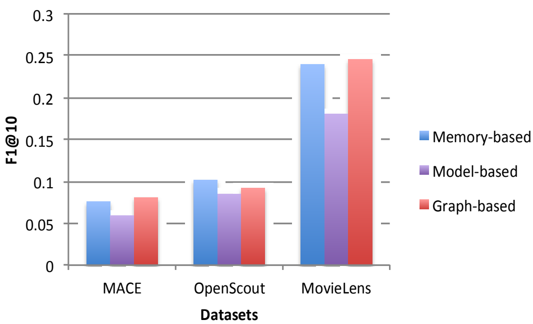



**Fig. 6.** F1 of the graph-based CF and the best performing baseline memory-based and model-based CFs, for all the datasets used

The graph-based approach collects recommendations for a target user from the neighbors reachable at a maximum path length (L). In the baseline kNN method, L always equals 1, which imposes the constraint of including only directly connected users to a target user. For the graph-based approach, choosing a value for L involves a trade-off. Increasing the L provides us with higher coverage but lower precision. Moreover, choosing larger path lengths can be more risky because of including malicious users in the recommendation procedure. Therefore, we set the maximum path length at L=3. Fig. 3 shows the F1 results of best performing memory-based CF (Jaccard kNN), model-based CF (BPRSLIM) compared to the graph-based CF. The horizontal axis (x) indicates the datasets used and the vertical axis (y) shows the values of F1.

As Fig. 3 shows, the graph-based approach performs best for MACE (8%) and MovieLens (24%) and the selected memory-based and model-based CFs stand in second and third place right after the graph-based CF. For OpenScout, the memory-based approach performs better with a difference of almost 1%. Note that the size of neighborhoods selected for Jaccard kNN on OpenScout was *k*=3, which is the smallest *k* (see Figure1) whereas *k* was set to 20 for the graph-based CF on other datasets. We did so because our strategy was to select the best-performing memory-based and model-based from steps 1 and 2 of this experiment. The Jaccard kNN performed best for *k*=3 in the case of OpenScout. If we consider the same *k* as we used for the graph-based (20), the Jaccard kNN's F1 (8%) is lower than the graph-based F1 (9.1%) for the OpenScout dataset. This shows that the graph-based approach performs well for all the datasets used.

In conclusion, according to the aggregated results presented in Fig.4, the graph-based approach can help to deal with the sparsity in the educational data coming from the social learning platforms. This is reflected by an improved F1, which is an effective combination of precision and recall of the recommendation made.

## 6 Practical Implications and Limitations

In the current study, it was difficult to make a comparison with the findings of related empirical research studies, such as the ones by Verbert et al. [2] and Manouselis et al. [10]. One of the reasons could be the use of different versions of the same dataset because the collected data belongs to different periods of time. For instance, for the MACE dataset, different versions are available. In fact, no unique version has been fixed for running the experiments, nor for making a comparison, in the community for Technology-Enhanced Learning (TEL) recommender system [13]. This problem originates from the already mentioned lack of a golden standard dataset in the educational domain, like the MovieLens dataset in the e-commerce world. In fact, it seems the TEL community, instead of aiming for their own single golden standard, should collect several representative datasets that can be used as a main set of references for different data studies on personalization and recommender systems. This observation was already made by the dataTEL project [13].

## 7 Conclusion and Further Work

The main goal of our study was to identify the most appropriate recommender algorithm that can support users to find useful resources in a social learning platform, such as the ODS platform. We conducted an offline data-driven study to evaluate a set of candidate recommender algorithms on a set of representative datasets similar to the ODS future dataset, as well as MovieLens from the ACM recommender systems community. We proposed a graph-based approach that aims to improve the process of neighborhood formation and thus, to improve the performance of baseline methods. The experimental study presented in this paper consists of three steps. First, we investigated which memory-based nearest neighbor methods best performs for the educational data used. Second, we evaluated state-of-the-art model-based methods using matrix factorization and Bayesian models. Ultimately, we evaluated the graph-based approach in comparison with the best-performing methods from the first and second steps of the experiment. The results showed that the graph-based approach outperforms baseline CFs and thus can tackle the sparsity problem in the data coming from social learning platforms. At present, we are working on using the matrix factorization (MF) methods in the graph-based approach, and to investigate whether this hybrid approach can improve the performance of each of the methods alone. The results presented in this paper serve as an initial step to investigate a recommender algorithm that can best fit the social learning platforms similar to the one used for ODS. As a further step, we intend to study usability of the selected recommender approaches by evaluating user satisfaction on novelty and diversity of the recommendations made.

**Acknowledgements.** This paper is part of a doctoral study funded by NELLL (the Netherlands Laboratory for Lifelong Learning) and the Open Discovery Space project. Open Discovery Space is funded by the European Union under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme for R&D. This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content. The work of Hendrik Drachsler has been supported by the EU project LinkedUp FP7 Program. Last but not least, we would like to thank the FIT Fraunhofer Institute, especially Dr. Martin Wolpers and Katja Niemann, for providing us with both the MACE and OpenScout datasets on a short notice. Without their support, this study would not have been possible.

## References

1. Vassileva, J.: Toward Social Learning Environments. Learning 1(4), 199–214 (2009)
2. Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., Duval, E.: Dataset-driven research for improving recommender systems for learning. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 44–53 (2011)
3. Golbeck, J.: Computing and applying trust in web-based social networks. University of Maryland at College Park, College Park (2005)

4. Massa, P., Avesani, P.: Trust-aware Recommender Systems. In: RecSys 2007 Proceedings of the 2007 ACM Conference on Recommender Systems, pp. 17–24 (2007)
5. Ma, H., King, I., Lyu, M.R.: Learning to recommend with social trust ensemble. In: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009, p. 203 (2009)
6. Fazeli, S., Zarghami, A., Dokoohaki, N., Matskin, M.: Mechanizing social trust-aware recommenders with T-index augmented trustworthiness. In: Katsikas, S., Lopez, J., Soriano, M. (eds.) TrustBus 2010. LNCS, vol. 6264, pp. 202–213. Springer, Heidelberg (2010)
7. Anjorin, M., Rodenhausen, T., Domínguez García, R., Rensing, C.: Exploiting Semantic Information for Graph-Based Recommendations of Learning Resources. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 9–22. Springer, Heidelberg (2012)
8. Manouselis, N., Drachsler, H., Verbert, K., Duval, E.: Recommender Systems for Learning, pp. 1–61. Springer, Heidelberg (2012)
9. Fazeli, S., Drachsler, H., Brouns, F., Sloep, P.: Towards a Social Trust-aware Recommender for Teachers. Spec. issue Recomm. Syst. Technol. Enhanc. Learn. Res. Trends Appl. (2013)
10. Manouselis, N., Vuorikari, R., Van Assche, F.: Collaborative recommendation of e-learning resources: an experimental investigation. J. Comput. Assist. Learn. 26(4), 227–242 (2010)
11. Cechinel, C., Sicilia, S., Sánchez-Alonso, M.-Á., García-Barriocanal, E.: Evaluating collaborative filtering recommendations inside large learning object repositories. Inf. Process. Manag. (2012)
12. Koukourikos, A., Stoitsis, G., Karampiperis, P.: Sentiment Analysis: A tool for Rating Attribution to Content in Recommender Systems. In: Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. (eds.) Proceedings of the 2nd Workshop on Recommender Systems in Technology Enhanced Learning 2012, 7th European Conference on Technology Enhanced Learning (EC-TEL 2012), pp. 61–70 (2012)
13. Drachsler, H., Bogers, T., Vuorikari, R., Verbert, K., Duval, E., Manouselis, N., Beham, G., Lindstaedt, S., Stern, H., Friedrich, M., Wolpers, M.: Issues and considerations regarding sharable data sets for recommender systems in technology enhanced learning. Procedia Comput. Sci. 1(2), 2849–2858 (2010)
14. Manouselis, N., Kyrgiazos, G., Stoitsis, G.: Revisiting the Multi-Criteria Recommender System of a Learning Portal. In: Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. (eds.) Proceedings of the 2nd Workshop on Recommender Systems in Technology Enhanced Learning 2012, 7th European Conference on Technology Enhanced Learning (EC-TEL 2012), pp. 35–48 (2012)
15. Thai-Nghe, N., Drumond, L., Grimberghe, A., Schmidt-Thieme, L., Krohn-Grimberghe, A.: Recommender system for predicting student performance. Procedia Comput. Sci. 1(2), 2811–2819 (2010)
16. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative Filtering Recommender Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) Adaptive Web 2007. LNCS, vol. 4321, pp. 291–324. Springer, Heidelberg (2007)
17. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: BPR: Bayesian Personalized Ranking from Implicit Feedback. In: UAI 2009 Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, pp. 452–461 (2009)
18. Rendle, S.: Context-Aware Ranking with Factorization Models. Stud. Comput. Intell. 330 (2011)

19. Ning, X., Karypis, G.: SLIM: Sparse Linear Methods for Top-N Recommender Systems. In: IEEE 11th International Conference on Data Mining, pp. 497–506 (2011)
20. Hirsch, J.E.: An index to quantify an individual's scientific research output. Proc. Natl. Acad. Sci. U.S.A. 102(16569) (2005)
21. Schmitz, H.-C., Scheffel, M., Friedrich, M., Jahn, M., Niemann, K., Wolpers, M.: CAMera for PLE. In: Cress, U., Dimitrova, V., Specht, M. (eds.) EC-TEL 2009. LNCS, vol. 5794, pp. 507–520. Springer, Heidelberg (2009)
22. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. 22(1), 5–53 (2004)

# Innovative Pedagogy at Massive Scale: Teaching and Learning in MOOCs

Rebecca Ferguson and Mike Sharples

Institute of Educational Technology, The Open University, UK
{rebecca.ferguson,mike.sharples}@open.ac.uk

**Abstract.** This paper looks at the implications for pedagogy of education at a massive scale. It begins by looking at educational approaches designed or adapted to be effective for large numbers of learners: direct instruction, networked learning, connectivism, supported open learning, and conversational learning at scale. It goes on to identify benefits and challenges of teaching and learning at scale. A grounded approach was used to analyse data from 18 MOOCs run on the UK-based FutureLearn platform. This identified benefits and challenges for learners, for educators and for society as a whole. These need to be addressed in two ways, through learning design and through platform design.

**Keywords:** MOOCs, pedagogy, massive scale, open learning.

## 1 Introduction

By definition, MOOCs are intended to offer effective online learning that is also open and massive. It has previously been difficult to achieve these three aims simultaneously: learning though mass public media is limited in its effectiveness, and successful large-scale online education is expensive to produce and deliver. A report for the Becta agency [1] on technology-assisted learning identified 17 essential learning practices that might be delivered or supported by technology. These include: expository, reflective, networked, tutored, case-based, problem-based, inquiry-driven, game-based, and constructive. Most have only been tested in computer-supported classrooms or in small-scale online courses. Which of these could succeed at massive scale, and can they be integrated into a coherent MOOC offering?

## 2 Literature Review

Literature relating to pedagogy at massive scale mainly comprises empirical accounts of attempts to design and run large-scale learning systems, ranging from a study of a mobile live video system for large-scale learning [2], to a data analysis of communications by 87,000 individuals on a MOOC [3]. The theoretical framing has come from discussion of the Network Effect [4] applied to learning. In brief, the Network Effect postulates that the value of a product or service increases with the number of people

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 98–111, 2014.
© Springer International Publishing Switzerland 2014

using it. For example, the telephone system becomes more valuable to users as more people are connected, offering more opportunity to make and receive calls. There are related disadvantages of large-scale networks, notably the effect of congestion from too many interconnected communications.

Also relevant is the effect of two-sided networks in which there are two distinct but inter-connected user groups. In the case of education, these would be educators and learners. Each needs to gain benefit from scale. Consider learners as points in a communicating network. As the number of learners increases, the number of opportunities to connect directly with other learners increases much more rapidly (with the square of the number of learners). The network becomes increasingly difficult to keep under central control. For a MOOC with 10,000 learners, there are 50 million ways in which pairs of them could connect directly. If the learning is to be successful, then ways must be found to manage this profusion of communication, to offer opportunities for direct teaching, mentoring and effective peer support.

The Network Effect was originally formulated to describe interconnections between computers. However, as Downes [5] points out, humans engage in meaningful communication, not just in data connection. There is value to be gained from the semantic relevance of these communications. Downes also indicates that a best set of connections cannot be determined in advance of the network being activated – it will depend on the abilities and contexts of the participants as they are revealed while a course is in progress.

The effects of scaling can also be explored in relation to pedagogy. Which teaching and learning methods are not sensitive to scale, and which may benefit from the effects of large-scale participation? There is over 70 years of evidence that exposition and instruction through educational radio and television can bring learning benefits to millions of people [6]. However, as Tyler indicates, “no broadcast is entirely self-sufficient” and its value depends on how it is used in context. For example, in the case of interactive classroom radio [7], a radio instructor provides lessons, prompting responses from the radio audience, and providing pauses for audience participation. Typically, a classroom teacher mediates this learner participation. The equivalent in MOOCs is to provide teaching by video, with opportunities for learners to pause the presentation and reflect or respond vocally, and for educators to embed formative quizzes into the video. The problem is that, while direct instruction by lecturing works at massive scale to engage learners and introduce big topics, this is a less effective approach to teaching than more active and collaborative pedagogies [8].

By contrast with direct instruction, networked learning is a process of collaborative meaning making through mutual support and interaction amongst learners [9,10]. This approach is also scalable to mass online participation. Sites such as Stack Exchange, with 4.9 million users[1], are networked communities that share expertise through organised clusters of questions and answers. Automated methods of reputation management and reward, through “liking” of postings and “badging” of people who have made valuable contributions, add a layer of community-defined value to the contributions. An example of a successful large-scale community that combines networked learning with reputation

[1] http://stackexchange.com/about

management is iSpot[2]. Over 40,000 people make observations of nature – including birds, mammals, plants and fungi – and post photographs online, with a brief suggested identification where possible. Other members of the community respond with further detail and identification. A reputation management system rewards with badges those who make many observations, have observations confirmed by others with high reputation, and provide accurate identifications [11].

Networked learning, and its close relation, connectivist learning, formed the educational basis of the original MOOCs developed by Siemens and Downes [12]. Connectivism views learning as distributed within a network, social, technologically enhanced and associated with recognizing and interpreting patterns. Knowledge transfer is achieved by adding nodes and thus expanding the network [13]. While these pioneered the MOOC concept, connectivist MOOCs require learners to direct their own learning and to have a high level of critical literacy in order to navigate web resources and engage with peers [14].

There is a fourth approach to online learning that has been shown to be sustainable at large scale. This is the "supported open learning" model adopted by The Open University[3], whereby teaching materials and assessment are delivered online, allowing students to work at times and locations convenient to them. A network of local tutors provides additional support including marking assignments, giving feedback, and offering help to students. Over 200,000 students at The Open University learn by this method, but it is costly to recruit, train and employ tutorial staff, so the approach is only viable on paid-for courses.

Therefore, a central challenge for massive free and open online courses is to develop innovative pedagogy that provides: the power of rich interactive media; benefits of active, reflective and collaborative learning; incentives through management of reputation; and opportunities for tutorial intervention and guidance – all capable of being deployed to tens of thousands of learners on a course, sustained over time. Other pedagogies, such as case-based, problem-based, inquiry-driven, and game-based learning may work at massive scale, but this is yet to be proven in practice.[4]

## 3 FutureLearn – Designing for Innovative Large-Scale Learning

One organisation seeking to develop a pedagogy that works at massive scale is FutureLearn. This is a company owned by The Open University that is currently in partnership with 36 universities, the British Museum, the British Library and the

[2] www.ispotnature.org

[3] http://www.open.ac.uk/about/main/the-ou-explained/teaching-and-learning-the-ou

[4] Arguably, game-based learning has been shown to scale through multiplayer "edutainment" games such as SimCity, though the launch of the online version of that game was beset with problems, see http://www.bbc.co.uk/news/technology-25715010

British Council, to deliver free online courses. The company has developed a new MOOC platform based on scalable web technology.

The FutureLearn management team made a decision early in the formation of the company to develop a MOOC platform and proposition that would support a social-constructivist pedagogy, based on the Conversational Framework [15,16]. In brief, this is a general theory of effective learning through conversations, with oneself and others, about the immediate world and about abstract concepts. To engage in successful conversations, all parties need access to a shared representation of the subject matter as well as tools for commenting, responding and reflecting.

The advantages of this framework include: it is derived from a theory of learning rather than instruction, it was devised originally to include interactions with and through technology, including large-scale "pervasive media" [17], and it embraces direct instruction and networked learning, as well as other types of learning, such as reflection and inquiry. Developing a MOOC platform around this framework led to some specific design decisions, outlined below.

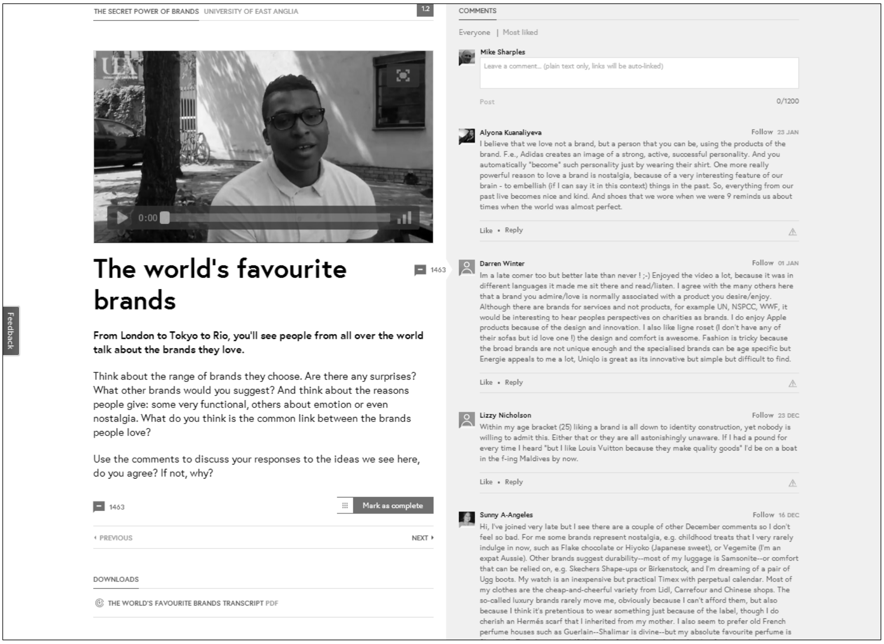



**Fig. 1.** Typical FutureLearn learning element ("step") showing some associated comments and replies. The flag indicates that 1463 comments have been added to the step.

Each teaching element (or "step") is associated with a free-flowing discussion, which is intended to emulate a "water-cooler conversation" about the immediate content (see Fig. 1). Each learner can see contributions from other learners and add a brief comment or reply. In terms of the Pask-Laurillard conversational framework, this is a conversation at the Level of Actions where discussion is directed towards interpreting the learning materials. Some FutureLearn content elements have attracted over 9,000 comments, so the approach will not scale if learners feel overwhelmed or obliged to read all the contributions. For this reason, the discussion is initially hidden and only revealed by clicking a button; it is structured as a simple flow with only one level of embedded replies; and learners can filter contributions to those that are most "liked" by others.

A distinctively different type of conversation can occur at the Level of Descriptions where the discussion explores differing perspectives on a more abstract topic or claim. This is implemented in FutureLearn as a Discussion Step and will also be supported through structured small-group discussions. These are designed to scale around those learners who are available online, through a "discussion bus" approach. As learners select the "small group discussion step" they are added to the discussion. Other learners join until the "discussion bus" (with room for around 20 people) is full. The discussion can start before the group is full. In these structured discussions, learners respond to a claim or question and are expected to read all the other contributions before replying. This approach should benefit from scale: with a larger pool of learners, there will be more people online at any time, so each "discussion bus" will fill more quickly. There will also be a wider variety of perspectives, from differing cultures, backgrounds, and abilities. In FutureLearn, peer review is offered as a formative learning activity, rather than as a method of assessment. This has also been designed to improve with scale. A learner is given a short, structured assignment. When the assignment is submitted, it is added to a pool and the submitter is sent an assignment that has not yet been reviewed. When the review is submitted, the writer of the assignment is alerted by email and can immediately read the review. The reviewer is offered an opportunity to read another assignment and submit a further review. The more learners there are online, the shorter the average wait for a review. Data from FutureLearn courses show a mean wait time of one to two hours (depending on the course), from submission of an assignment to receiving a first review.

## 4 Implications of Scale

Media discussion and promotional hype around MOOCs have focused on "massive" in terms of the very large numbers of students enrolled on these courses. "Massive" is presented in terms of potential economic benefits – huge numbers of people can be exposed to the same curated set of high quality resources and can gain access to teaching materials assembled by globally respected universities. Fewer trained educators are, potentially, required in order to educate greater numbers of learners.

In this section, we examine some of the other implications of massive for both learners and educators, as well as for MOOC designers and for society as a whole. In order to do this, we examine data from the 18 courses that started their first presentation on FutureLearn between 14 October 2013, when the first MOOC was offered on the platform, and the date of writing, 3 February 2014. Fourteen different UK universities were responsible for these courses. The data that were analysed consist of the course materials, course emails and course discussions that were available to anyone who enrolled in these open online courses, as well as associated activity in social media.

The terms and conditions of FutureLearn state:

*7.11 Any Learner Content that is published on the public discussion areas of the Website (for example, the forums or areas where posts are made) will be subject to a Creative Commons Licence (Attribution-Non Commercial-NoDerivs; BY-NC-ND).*

For this reason, the authors of all direct quotations used within this paper are named and are not anonymised or assigned pseudonyms.

This paper addresses the research question

*What are the implications for pedagogy of teaching and learning at massive scale?*

In order to identify emergent themes within the data, we adopted a grounded approach to analysis. The intention was not to offer a comprehensive analysis of these MOOCs, but to address our research question using a large dataset including contributions from many thousands of people who were experiencing massive online teaching and learning for the first time. Some of the themes identified here were found many times within the dataset; some only once or on only one course, but each has the potential to impact on a pedagogy of teaching and learning at scale. The themes that emerged can be grouped under two headings: "Advantages of massive pedagogy" and "Challenges of massive pedagogy". The following sections explore these themes in more detail and provide illustrative examples from the dataset.

## 5 Advantages of Massive Pedagogy

Literature relating to education identifies many advantages to learning together, including collaboration, cooperation and co-construction of knowledge. However, these advantages are not necessarily scalable; cooperating with ten thousand people is unlikely to be better than cooperating with ten. In addition, the pedagogic underpinnings for learning design are rarely made explicit within course material, and so these benefits of learning together are not foregrounded within the dataset. Grounded data analysis brings to the fore a set of advantages associated with engaging thousands of learners from all over the world. These can broadly be grouped as: advantages for learners, advantages for educators and advantages for society.

### 5.1 Advantages for Learners

Massive participation gives learners access to support from a wide range of other learners, to resources provided by those learners in the form of discussion and links, and to a range of diverse cultural perspectives.

As MOOC learners can outnumber educators by 1,000 to one or even more, the importance of other learners in providing support is clear. Tim Lenton, lead educator on the "Climate Change" MOOC emailed learners early on:

*Even before the course begins you can get in touch with your fellow students by using the hashtag #exClimate on Twitter – start the conversation about climate change now!*

And, as soon as the course began, he emailed:

*I will be encouraging you to take part in online discussions with your fellow learners as the course progresses. Use these discussions to ask questions if you're struggling, and help other people if they're stuck.*

These discussions were not confined to the FutureLearn platform. Online, learners took the initiative and set up Facebook groups to discuss courses, including "Richard III" and "Forensic Psychology". Teams from dental practices signed up together to work through "Dental Photography" and groups of teachers signed up to work through a MOOC together "to use the resource in staff meetings and as part of their professional development" [Helena Gillespie, Teaching Computing].

As well as educators, many learners were positive about the opportunities for discussion. Postings on the last day of the "Ecosystems" course included:

*Thanks to David and team for a very interesting and enjoyable introduction to a subject I will certainly be pursuing on OpenLearn and elsewhere. Thanks also to fellow students for some very informative and helpful discussions. [Tom Strathdee]*

*Surprised to find that I really enjoyed the discussion board; some great postings and links. Also good to have your position and understanding challenged. Thanks to all! [Andrew Revell]*

*Thanks also for all the links to interesting websites, both from the course team and from fellow students. [Julia McDougall]*

Within MOOCs, learners provide and swap perspectives on content that are informed by their experience and their cultural background. A discussion of brand names on the "Secret Life of Brands" was enriched by contributions from different countries.

*To me the ultimate brand recognition is when the brand name has actually become the name of the product like Hoover for vacuum cleaner, Frigidaire in French for a fridge. [Delphine Lemaire]*

*one of the strongest brands in the UK is "Orange." However, people still go to buy a "mobile Phone" and not an Orange! (er, if you see what I mean). Talking of Orange, it is a great example of localised branding messaging. "The Future's Bright, The Future's Orange" was their message for many years in the UK - but NOT in Northern Ireland where that would have alienated a huge section of the community! [Joss Sanglier]*

*In Ukraine they call pampers all disposable diapers as it was the first brand in the market and it still stays the strongest. [Natalya Dubois]*

In some cases, the perspective from another country could cast a new light on an entire subject. A discussion on "the potential uses of dental photography in your practice" focused mainly on aspects such as records, analysis, treatment, planning, practice and legal concerns. A post from Paraguay opened up another range of possibilities with the potential to expand dental photographers' view of the role and importance of their specialised skill:

*I'm one of the very few photographers in an area trice as big as Scotland (the Paraguayan Chaco), where 8 different cultures live together. There's no record-keeping or proper research done with regards of dental treatment, specially among the most remote areas on the country. I believe there's huge opportunities to enhance the dental service through the powerful medium that photography provides, by providing material for research and education. Mainly to inform central and local government and other decision makers, about the realities the region faces, but specially to educate patients (elderly and children) dental workers and other professionals, that sometimes have to travel hundreds of kilometres on very bad roads, in order to reach a small town with no proper facilities but where people in need of proper treatment live. [Elias Adan Gimenez Feliu]*

### 5.2 Advantages for Educators

Educators identified affective benefits of teaching in MOOCs; they also saw opportunities for increased access to resources, and a motivation to develop their teaching practices.

Educators' blogs and weekly emails to enrolled learners provided opportunities for them to express their feelings about teaching such large groups. A sense of pride and excitement emerges from these messages, and a sense of joy that it is possible to share personal passion and enthusiasm with learners worldwide.

*I am very passionate about the study of language based on naturally occurring speech and writing. So getting more people to know about it and be able to do it is my goal! [Tony McEnery, Corpus Linguistics]*

*Here in Stratford we're immensely excited to see how the discussions are shaping up, and to welcome such a range of online learners from all around the world […] It is very exciting to have such a rich and diverse online community thinking about Hamlet and comparing notes about it with one another and with us. [Michael Dobson, Hamlet]*

*I have learnt that people all over the world have engaged with the work. I've seen discussion responses from places such as India, Indonesia, USA - that has been very exciting, to see such world-wide participation. [Harriet Jones, Preparing for Uni]*

The scale of interest in MOOCs means that educators and others involved in developing courses sometimes have opportunities to develop bigger budget resources than would otherwise be possible. For example, the team responsible for "Forensic Science" was able to make a mini-series of a case study, bringing the techniques under discussion to life with a narrative that kept learners engaged right up to the final cliff-hanger verdict on the last evening of the course. Resources from external providers were used to enrich courses, including readings to support lectures, "donated by publishers including Bloomsbury, Edinburgh University Press and Routledge" [Tony McEnery, Corpus Linguistics]. Other courses were able to include interviews and discussions with external experts. “Hamlet” included interviews with Jonathan Slinger and Pippa Nixon from the Royal Shakespeare Company. “Corpus Linguistics” included conversation sessions with leading academics. Course lead Tony McEnery commented, “What about the ‘in conversations’? I must now confess a guilty secret – I have really, really, really enjoyed doing these”. The size of MOOCs mean some are also able to attract sponsorship and collaboration – the “Teaching Programming” course, which relates to a recent change in England’s National Curriculum for schools, was supported by both British Telecom and by the Computing at School organisation.

Learning wholly online is a new experience for most people, and a massive cohort offers the possibility of sharing learning skills and experience of higher education, as well as subject knowledge, with large numbers of learners.

*I have tried to make sure that the experience you have is as close to being at University, in spirit, as possible. The University is a place of thought, discussion and lecture. That is what, I hope you will agree, the MOOC is when you start to study it next year. [Tony McEnery, Corpus Linguistics pre-course email]*

Much of this is implicit in educators’ work; but in some cases it is foregrounded. “Fairness and Nature” included study skills as a part of the course, focusing on preparing to learn, listening and reflecting, making notes and communicating with others. Lead academic Tim Lenton’s Climate Change blog begins, “Practise what you preach, as they say. So I have decided to do some reflective learning and blog about my experience with our Climate Change MOOC.” In many cases, this was a two-way exchange of ideas, with educators not only sharing tips about how to learn in MOOCs but also receiving feedback about subject matter and approaches to teaching that work well in these environments. Such feedback was sometimes explicit, in the form of comments, and at other times implicit, in the form of engagement and withdrawal. It could also take place outside the platform and before or after the course start. One pre-course email stated,

*Looking at those of you who follow me on Twitter has proved to be very influential in shaping the course – we really want to provide a course on which you will find some real value. [Tony McEnery, Corpus Linguistics]*

## 5.3 Advantages for Society

The extended reach of MOOCs offers the potential for far-reaching effects. These courses can be used to develop tools and resources for use in other contexts; they can be used to change professional practice, can increase access to education and are seen to have the potential for global impact.

When FutureLearn was launched in 2013, the UK Universities and Science Minister, David Willets, said, “I encourage all our institutions to explore the opportunities offered by new modes of technology, such as MOOCs. This will keep the UK ahead in the global race to deliver education in worldwide markets.”[5]

---

[5] `http://www.bbc.co.uk/news/business-24109190`

Educators leading the MOOCs also have an interest in teaching and learning with worldwide impact:

*I am particularly keen for the MOOC to reach audiences that don't have access to the learning experience we provide here at Leeds – such as those in Nepal, which features prominently as one of our case studies. [Jon Lovett, Fairness and Nature]*

In the case of professionally focused MOOCs, such as "Dental Photography" and "Teaching Programming", courses can provide continuing professional development (CPD) for many thousands of professionals in a short period of time. These training opportunities can extend beyond course boundaries as learners set up local groups and social media groups to continue conversations and build on what they have learned. The University of Southampton already has a LinkedIn group that is designed to keep MOOC alumni in touch with each other after the end of a course.

The resources created for and within MOOCs can also be used and repurposed outside the course environment, and some course teams encourage this to take place:

*You are free to repurpose its content as long as you follow the licence conditions. You don't need to ask our permission, although we would like to know! And our Flickr site, Scoop it site and Delicious site will all stay open. [Sarah Speight, Sustainability, end-of-course email]*

"Ecosystems" students contributed hundreds of images to the iSpot nature identification site, while "Dental Photography" students created a shared resource of images on Flickr. Extensive user testing at scale of the AntConc concordancing software used on the "Corpus Linguistics" MOOC enabled the development of a new version of this freeware tool, thus upgrading the toolset available to everyone working in this field.

## 6 Challenges of Massive Pedagogy

In many cases, the challenges of teaching and learning are the flipside of the advantages. In an environment with tens of thousands of people, where a discussion may contain over 9,000 contributions, both learners and educators face the challenge of navigating and filtering resources. Learners gain access to support from a wide range of other learners, but they need ways to ensure that they receive good quality, well-informed support. They can see a wide range of resources and perspectives but, at the same time, require ways in which to filter and make sense of these resources.

The flowing discussions alongside the course materials in FutureLearn offer an easy way to see and add to the immediate responses of other learners, but some learners report being disorientated by the large number of contributions. A similar problem of disorientation was reported in the early days of the worldwide web. The "lost in hyperspace" phenomenon "can refer to any of the following conditions: users cannot identify where they are; users cannot return to previously visited information; users cannot go to information believed to exist; users cannot remember what they have covered; and users cannot remember the key points covered" [18].

When MOOCs are regarded as an innovative approach to teaching and learning they can inspire and excite educators, but educators need ways to maintain these positive feelings when the courses repeat and the novelty wears off. Increased access to resources requires time both to locate and embed these resources. Developing effective teaching practice in new environments is challenging to do in full view of an audience of thousands, and requires time and effort that may not be accounted for in institutions' workload planning.

Many of these challenges need to be addressed at platform and institutional level. Others relate to accessibility, particularly in relation to age, location and disability. Many MOOCs act as university showcases, designed to attract potential students. This means that a large number of MOOC learners are teenagers, which raises issues related to suitable content and to child protection when taking part in online activities. The "Forensic Science" course posted a warning each week:

*Warning: Some of content presented in this program may be distressing to individuals, particularly younger learners. Notwithstanding, the material is representative of that encountered by forensic scientists and we have presented it in an objective and professional manner.*

"Forensic Science" encountered a different set of accessibility problems when using Google Hangouts to support live interaction on the course.

*I'm not sure I can participate live -- the timing is difficult (time difference I'm in Canada and will be working. [Jean Read, Forensic Science]*

In theory, those unable to attend live should have been able to replay the event on YouTube but, once again, location denied access to some participants

*Not able to click YouTube in China. What a pity ! [Yue Shi, Forensic Science]*

Disability also poses accessibility challenges at scale. Start-of-course and end-of-course surveys on one FutureLearn course received responses from around 10% of learners. These surveys suggest that 14% of learners had some form of disability. Subdivisions of this dataset are small, so may be misleading, but 0.4% of respondents to the start-of-course survey stated that they were blind, and more than 1% stated that they were dyslexics. On a course with 10,000 learners, this would equate to 40 blind learners and 130 dyslexic learners on every presentation of the course. This raises the challenge of making courses and their assessment accessible to all learners.

## 7 Discussion

Analysis of these 18 courses shows that massive pedagogy requires more than an underpinning framework. Conversational learning can and does scale; the FutureLearn discussion forums are bustling with activity, and around a third of learners are actively contributing to these discussions, alongside those who are benefiting by reading and reflecting on these contributions. However, until all the benefits and challenges of learning at scale are actively addressed, there will always be room for improvement on any MOOC platform. Addressing these elements requires a twin-pronged approach, planning for them in both the learning design and the platform design.

A course with thousands of registered learners has the potential to offer them access to support from a wide range of peers. In order to do this, its learning design needs to take into account the points at which learners are likely to need support; needs to build in opportunities for asking questions, raising concerns and asking for help, and needs to build in motivation for offering help to others. This will not be possible unless the MOOC platform not only provides opportunities for learner communication, but also includes ways of judging which people are offering helpful and reliable advice. Such judgments cannot be made without contextual information. "Online we use social factors (rating and voting) to assess reliability, the user's profile and badges to demonstrate competence" [19] and these are all tools which the platform can provide to support an effective learning design.

A massive course can go far beyond any traditional course in providing access to a huge range of resources and a global range of perspectives provided by those registered on the course. Again, the learning design needs to build in opportunities to make use of these benefits. MOOC learners are often fitting their learning in around many other activities; they need resources and perspectives that will extend their understanding of the course, rather than ones that distract their attention. When will these resources and perspectives be solicited, and how will learners use them? The platform design has a role to play here. Likes, ratings, tags and analytics can help both learners and educators to sort through contributions and to locate the ones that will be helpful for their learning or teaching.

For educators, MOOCs can offer affective benefits, but initial excitement will wear off if there is nothing to sustain it. Learning design has a role to play here as well; it is possible to build in points where learners share stories of success or raise difficult problems that the network of learners can work to solve together. In terms of platform design, there is no need to leave educators siloed in their individual MOOCs. Forums, discussion areas and chances to meet up both online and offline would offer opportunities to extend and share practice, to share possibilities for creating or accessing resources that are not available when working at a smaller scale, and to build on success. These areas would also offer opportunities for another benefit of massive – the impetus to develop new teaching practices.

In terms of society, one of the great benefits of MOOCs is the opportunity to open access to higher education to people who would be excluded from study at traditional universities. In order to achieve this goal, the learning design needs to be constructed so that none of its elements unnecessarily exclude people on the grounds of disability, age or location. From a platform perspective, attention to detail – for example, always including alternative text for pictures and ensuring that PDFs can be read in full by screen readers – will help to keep MOOCs not only massive but also open. More broadly still, massive education offers the opportunity to address "wicked problems" – like climate change, which cannot be resolved by one individual, one organisation or even one country. From a learning design perspective, this involves thinking big when setting the learning outcomes for a course. From a platform perspective, this involves building in opportunities to scale even further, so that a variety of courses from different universities can deal with aspects of the same problem, can develop learning outcomes together and have opportunities to share and pool the knowledge that they generate.

## 8 Conclusion

This paper has looked at the implications for pedagogy of learning at scale. With MOOCs now regularly attracting tens of thousands of learners, it is important that educators are able to make use of forms of teaching and assessment that benefit from high student numbers. Four approaches that were in use before the emergence of MOOCs, and that have been designed or adapted to scale are direct instruction, networked learning, connectivist learning and supported open learning. However, each of these raises challenges. Direct instruction is not as effective as other approaches to teaching. Networked learning can be overwhelming when learners have to deal with too many interconnected communications. Connectivist approaches require learners to direct their own learning and to possess a high level of critical literacy. Supported open learning is an expensive option on courses that are provided free of charge. A fifth approach is being developed specifically for MOOCs, exemplified by FutureLearn, and is based on a well-established theory of effective learning through conversations with oneself and others. Other pedagogies, such as game-based, problem-based, and inquiry-led learning still need to be adapted and demonstrated for massive scale.

Analysis of MOOCs on the FutureLearn platform has identified advantages and challenges of teaching and learning at scale, which should be taken into account in learning design and from a platform perspective. For learners, scale offers access to support from a wide range of other learners, to resources provided by those learners, and to a range of perspectives. For educators, scale offers affective benefits, opportunities for increased access to resources, and a motivation to develop teaching practice. For society, scale offers potential to develop tools and resources for use in other contexts, to change professional practice, to increase access to education and to achieve global impact. The challenges of scale include the need to navigate, filter and make sense of resources, and for learners to be able to access good quality, trustworthy support. MOOCs offer the potential to open up education for those who were previously excluded but, in order to do so, must take on the challenges associated with disability and disadvantage.

## References

1. Manches, A., Phillips, B., Crook, C., Chowcat, I., Sharples, M.: Year 3 Final Report: Shaping Contexts To Realise the Potential of Technologies To Support Learning. Curriculum and Pedagogy in Technology Assisted Learning project. Report by University of Nottingham and Sero Consulting Ltd. working in association with Becta (2010), `http://www.naace.co.uk/capital` (accessed online February 21, 2014)

2. Ullrich, C., Shen, R., Tong, R., Tan, X.: A mobile live video learning system for large-scale learning – system design and evaluation. IEEE Transactions on Learning Technologies 3(1), 6–17 (2010)
3. Gillani, N.: Learner communications in massive open online courses (OxCHEPS Occasional Paper No. 53, University of Oxford (2013), http://oxcheps.new.ox.ac.uk/MainSite%20pages/Resources/OxCHEPS_OP53.pdf (accessed online February 20, 2014)
4. Metcalfe, R.M.: It's all in your head. Forbes (April 20, 2007), http://www.forbes.com/forbes/2007/0507/052.html (accessed online February 20, 2014)
5. Downes, S.: The Personal Network Effect. Blog posting (2007), http://halfanhour.blogspot.co.uk/2007/11/personal-network-effect.html (November 4, 2007)
6. Tyler, K.: Recent developments in radio education. The English Journal 28(3), 193–199 (1939)
7. Imhoof, M.: Interactive radio in the classroom: ten years of proven success. Development Communication Report 48, 4–5 (1985)
8. Prince, M.: Does active learning work? A review of the research. Journal of Engineering Education 93(3), 223–231 (2004)
9. Illich, I.: Deschooling Society. Calder and Boyars (1971)
10. Goodyear, P., Banks, S., Hodgson, V., McConnell, D. (eds.): Advances in Research on Networked Learning. Kluwer Academic Publishers, London (2004)
11. Clow, D., Makriyannis, E.: iSpot analysed: participatory learning and reputation. In: Proceedings of LAK 2011, pp. 34–43. ACM, New York (2011)
12. Downes, S.: Connectivism and Connective Knowledge (2012), http://www.downes.ca/files/books/Connective_Knowledge-19May2012.pdf (Book published online)
13. Siemens, G.: Orientation: Sensemaking and Wayfinding in Complex Distributed Online Information Environments. PhD thesis, University of Aberdeen, Aberdeen, UK (2011)
14. Kop, R.: The challenges to connectivist learning on open online networks: learning experiences during a massive open online course. International Review of Research in Open and Distance Learning 12(3), 19–37 (2011)
15. Pask, G.: Conversation Theory: Applications in Education and Epistemology. Elsevier, Amsterdam (1976)
16. Laurillard, D.: Rethinking University Teaching: A Framework for the Effective Use of Learning Technologies, 2nd edn. Routledge Falmer, London (2002)
17. Pask, G.: Minds and media in education and entertainment: some theoretical comments illustrated by the design and operation of a system for exteriorizing and manipulating individual theses. In: Trappl, R., Pask, G. (eds.) Progress in Cybernetics and Systems Research IV, pp. 38–50. Wiley (1975)
18. Thimbleby, H., Jones, M., Theng, Y.L.: Is 'lost in hyperspace' lost in controversy? Proceedings of the Eighth ACM Conference on Hypertext, Posters and Demonstrations (1997), http://www.cs.waikato.ac.nz/oldcontent/mattj/acmhypertext.pdf
19. Liu, H., Macintyre, R., Ferguson, R.: Exploring qualitative analytics for e-mentoring relationships building in an online social learning environment. In: Proceedings of LAK 2012, ACM, New York (2012)

# Mobile Inquiry-Based Learning with Sensor-Data in the School: Effects on Student Motivation

Olga Firssova, Marco Kalz, Dirk Börner, Fleur Prinsen, Ellen Rusman, Stefaan Ternier, and Marcus Specht

Welten Institute - Research Centre for Learning, Teaching and Technology
Open University of the Netherlands
Heerlen, The Netherlands
{olga.firssova,marco.kalz,dirk.boerner,ellen.rusman, fleur.prinsen,stefaan.ternier,marcus.specht}@ou.nl

**Abstract.** The paper discusses the design, implementation and evaluation of a pilot project that integrated inquiry-based learning with mobile game design and introduced mobile devices and sensors into classroom learning. A 5-week classroom inquiry learning project on energy consumption was designed and implemented as a mobile serious game. While engaging in the game and training inquiry skills, students were introduced to basic concepts in the energy domain and to everyday practices of energy consumption in their direct environment. The design was based on a model of inquiry-based science learning with social and open mobile tools developed in the European research project weSPOT. The pilot indicated that using an ubiquitous tool does not necessarily help sustain student motivation. There are indications of gender differences in motivation related to use of mobile devices for learning. These differences should be taken into account in the design of activities. Implementation of the inquiry-based learning model in conjunction with a mobile game scenario into the school practice confirmed the importance of good design with sufficient testing and teacher ownership.

**Keywords:** mobile learning, motivation, inquiry-based learning, environmental literacy, energy consumption, design, implementation.

## 1 Introduction

In the last ten years social and mobile media have become an important, if not a primary communication channel for young people and a stable part of their daily life and daily routines. The potential of mobile media for learning is not disputed and a body of knowledge on affordances of mobile devices as learning support tools keeps growing [1,2,3]. However, instantiations of mobile learning in school practice are still scarce in European countries. Teachers do not feel comfortable using these technologies and mobile devices remain more often banned from classrooms than used for learning purposes [4,5].

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 112–124, 2014.
© Springer International Publishing Switzerland 2014

Discussing the added value of mobile learning, Ally [6] emphasizes easy access to knowledge sources through ubiquitous and cheap portable devices. From the active learning perspective, interaction with artefacts and persons in and across learning contexts, is important. Smartphones and tablet computers with enhanced facilities for instant and asynchronous messaging support such interaction and make knowledge exchange and knowledge building possible in any-time, any-place modes in the most direct sense of the word. Curiosity-fed questions arising at the spur of the moment at a fieldtrip, a museum visit or a workplace situation can be captured by a single click, annotated in text or voice and directly saved and uploaded into a personal school learning space or a space shared with others. Thus, mobile devices can create bridges between formal and informal learning, bringing the richness of authentic learning experiences into the classroom [7,8].

The topics of sustainable development, renewable energy sources, energy efficiency, energy consumption and conservation belong to urgent societal problems in need of solutions and are included in various national formal science curricula both in primary and secondary education [9,10]. Not much, however, is known about how school knowledge on issues such as energy consumption, environmental protection, sustainable development or energy conservation interact with daily life experiences and views of young people on these issues, whether connections between school subjects, daily life experiences and global societal issues help make abstract school knowledge meaningful and usable in practice [11,12].

Affordances for capturing moments of curiosity that mobile devices have, can help to increase awareness of what is happening in the surrounding world and contextualize learning. These affordances make such devices natural partners of environmental education [13,14]. For instance, sense-making and deep learning can be enhanced when learners consult background sources directly during fieldtrip activities and communicate on these points with experts or fellow students elsewhere without leaving the field. Thus, mobile devices become effective tools during exploratory activities outdoors and during the process of sense-making of these activities [15]. As demonstrated by Börner and colleagues [16,17], a serious location-based mobile game and ambient displays can help raising awareness of the energy consumption patterns at the workplace and can lead to behavior change.

Implementing pedagogical approaches that focus on the links between formal school learning and everyday life experiences, such as inquiry learning, has been known to present challenges for curriculum designers and teachers [18]. Edelson, Gordon and Pea [19] indicate that student learning will be undermined in case of failure to address several challenges, including that of sustaining the learner's motivation throughout the inquiry process. Other challenges are addressing gaps in background knowledge, scientific investigation skills (including data collection and analysis techniques) and skills in project management. These authors also emphasize the necessity to address the practical constraints of the learning context in design and implementation of inquiry-based learning [p.399].

Available research, and innovative approaches of the early adopters of new technologies [cf.,20,21,22], indicate that ubiquitous tools such as mobile devices can contribute to successfully tackling some of these challenges. There is evidence that using mobile devices in the context of gaming is highly motivating for learners [23]. Further tests in ecologically valid situations of average technology users and authentic school settings remain, however, highly relevant for understanding what works and why for optimal use of these ubiquitous tools as learning enablers.

To support understanding of how an inquiry-based learning activity in the form of a serious mobile game can be integrated in regular school practice and support inquiry-based science learning, a game on the topic of energy consumption was conducted. Students were triggered to investigate and compare energy consumption of different appliances in the school building, test assumptions on energy efficiency and present results to each other, teachers and management. A theoretical model of inquiry-based science learning with mobile and social media, the weSPOT[1] IBL Model, formed the framework of the designed learning activity. ARLearn, a toolkit for designing mobile and location-based learning games [24] was used to design and deliver instructional scripts to guide inquiries and provide just-in-time access to relevant information to the players. Experiences from Mindergie, a mobile game on energy consumption awareness at the workplace [17], were used to implement a sensor network in the school building, to provide input for inquiry activities and to act as triggers for learning.

The study focussed on testing feasibility of integrating such an activity and new technologies as mobile devices and sensors in classroom learning. It also aimed at evaluating whether students were and remained motivated to learn and conduct mobile-led activities at school. Furthermore, as a first implementation of a new pedagogical approach in a particular school setting, the pilot was expected to shed light on the practical constraints of implementing inquiry-based learning with mobile devices in the school practice.

The study set-up and general findings were introduced by Kalz and colleagues [25]. The findings suggested that practical constraints, in the first place the defects of Internet network connections, hindered the pilot. Furthermore, effects on knowledge and motivation were manifested. Learners with a low level of prior knowledge turned to benefit most from the activity. As far as cognitive outcomes are concerned no gender differences were found. An unexpected outcome was a general decrease in motivation by all participants and in particularly by girls who were significantly more interested in participating in the activity at the start. To get a clearer understanding of the effects of the intervention on student motivation and formulate lessons for design, additional analyses were conducted. This paper elaborates on these results and the lessons learned from the pilot.

The next sections give an overview of the applied weSPOT IBL model, elaborate on the pilot and its results, including the pilot implementation and practical constraints. Finally, the authors reflect on the implementation of innovative technologies in relation to motivation issues based on the findings of the study.

---

[1] weSPOT Project - IST (FP7/2007-2013) under grant agreement N° 318499.

## 2 Inquiry-Based Science Learning to Foster Environmental Literacy: The weSPOT Model

An inquiry-based learning model developed in a European project weSPOT– Working Environment with Social and Personal Open Tools for Inquiry based learning was used as a framework for game design. Three basic assumptions form the backbone of the weSPOT inquiry-based learning framework:

- Everyday experiences feed natural curiosity of young learners and can enhance formal classroom learning in which knowledge of scientific concepts is developed.
- Personal experiences and insights can help to understand theoretical concepts taught in the classroom.
- Development of inquiry skills by doing inquiries should mirror the process of systematic scientific observation and experimentation as well as consistent and critical reasoning that are standard in scientific communities [26].

The weSPOT IBL model views inquiries as learning activities in which students develop knowledge and understandings of scientific ideas by actually doing research. It does not prescribe how learning should be designed but offers flexibility to teachers and learners in the way inquiries can be done. According to the weSPOT model, inquiry activities can be designed and led by teachers, by learners or by both. They range from highly structured to open inquiries offering a maximum of flexibility and freedom for learners to set inquiry goals to define inquiry results and organize the inquiry process [27].

Distinct features of the weSPOT inquiry model are the interrelatedness of phases of the inquiry cycle and the flexibility that it offers to teachers in choosing to focus either upon the complete inquiry cycle or upon one or several phases in organizing learning. The proposed phases mirror the broadly accepted division of the research process which encompasses the stages of setting the research goals and formulating research questions, operationalization, data collection, data analysis and interpretation and communication of results.

The weSPOT IBL model makes explicit links between the domain knowledge, domain-related and generic inquiry-related skills, school-based inquiry activities, everyday life experiences and the context in which knowledge and skills are applied [28]. Phenomena or developments around them trigger learners' curiosity and provoke asking questions that give input for new inquiries or help to deepen ongoing inquiries. To find answers, learners undertake various activities, including experimentation, data collection and information searches. While collecting data and communicating their findings in authentic contexts, learners (learn to) apply domain knowledge and both domain-related and generic inquiry-related skills that they master at school. Figure 1 provides a general overview of the inquiry process according to the weSPOT model.

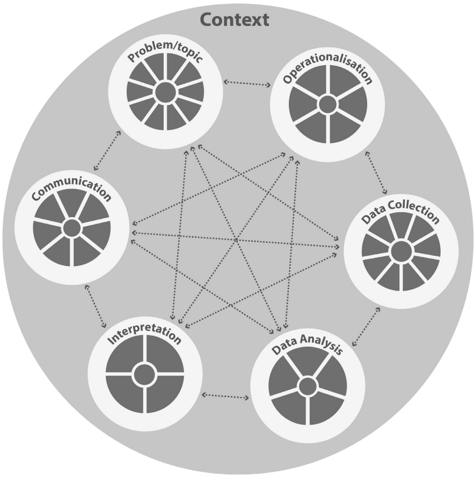



**Fig. 1.** weSPOT model and the constituent phases of the inquiry process. Each phase includes a range of tasks and activities [26].

Pilots with the users, i.e., with schoolteachers and students are being held from the early stages of the project. The reported pilot study tested a possible application of the weSPOT inquiry based learning model in combination with a mobile app prototype, which guided learners through the inquiry.

# 3 The Pilot

In the pilot:

- A series of instructional mini-scripts (games) was designed to guide students in individual and collaborative inquiry activities on the topic of energy consumption.
- These scripts were delivered through an app for personal mobile devices as games including game elements, such as information units, challenges, fun elements and incentives.
- A sensor-based system was used to collect energy consumption of appliances within the school building.
- Both prior to the pilot and after it, data was collected on learners' basic knowledge of energy consumption and on students' motivation for learning about energy consumption.

## 3.1 Method

**A. Participants**

The game was integrated in science lessons as a 5-week project on energy consumption. 13-year old K-8 level students of a secondary school in the Netherlands spent weekly between 2 and 4 hours a week on this project (n=85). 58% were girls. In their activities three science teachers guided students.

### B. Materials and Instruments

*Mobile games.* In the pilot 7 mobile games were designed to guide students in their inquiry activities. These games contained knowledge rich information units as text, video or links to web-based sources on the topic of study: energy sources, electrical power units of measurement, efficiency calculations, energy consumption; on phases in the inquiry process and on relevant methodological concepts as a research question, a hypothesis, data collection, presenting results, etc. The information units were presented by a young story teller. These units were combined with instructional prompts on tasks and challenges for the students as doing a short multiple choice quiz or finding a QR code that should be scanned to get access to further information or instructions. Figure 2 gives an illustration of the display on a task prompt, a quiz unit and an information unit with a test prompt (in Dutch).

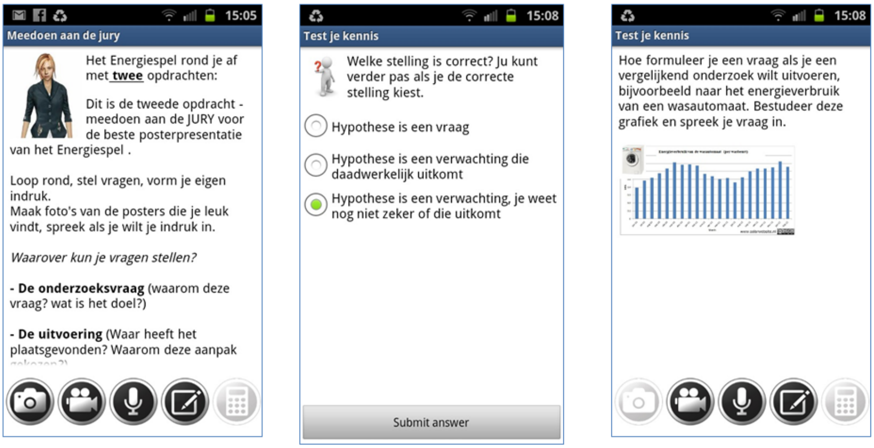



**Fig. 2.** Android device displays featuring exemplary ARLearn tasks

Some tasks became available only when certain activities were completed or particular quiz tasks were performed (i.e., data collection tasks were to be completed to proceed). Once started all tasks remained available.

*Plugwise sensor network.* Information on energy consumption of appliances within the school building was provided by the Plugwise system. Smart meter plugs that measure and store individual consumption data build up a wireless network that can be accessed using the bundled Plugwise software [29]. The system was set in the school building, in such a way that single appliances could be accessed. In total 50 appliances and power plugs throughout the school were equipped with smart meter plugs. This setup included three classrooms, a meeting room for teachers, some shared facilities (e.g. printer, vending machines, air conditioning), as well as some stations for experiment setups (e.g. for comparing the energy consumption of different light bulbs). Each appliance has been measured separately. In order to establish communication within the network, each plug had to be positioned within a range of 5 to 10 meters to another plug. The needed software has been installed on a laptop computer. Utilizing a USB dongle the software communicated directly with the network. To allow remote access to the system and the data, the web server functionality had been activated. A similar network set-up was earlier realized in the Mindergie game [17].

*Personal mobile devices.* According to an inventory held at the beginning of the school year 80% of the students had personal mobile devices with Internet access through Wi-Fi. Around 10% of these students had IOS supported devices (iPhones), the rest used Android devices. According to the anonymous evaluation inventory held after the pilot (n=81) 95% of all participants had a smartphone, with 16% (n=14) iPhones and the rest Android devices. 15 ipad mini's of the school were made available for students who had no smartphones.

*Classroom teaching materials.* Upon explicit teacher request in addition to the mobile scripts, 5 power point presentations on each phase of the inquiry cycle and a presentation on domain related concepts as energy sources and energy efficiency were designed for classroom use to provide an alternative information source for the students who did not have mobile devices. Duration of the presentations varied from 15 to 30 minutes. Each presentation highlighted a specific phase of the scientific inquiry cycle and included the content of the weekly inquiry tasks. These presentations contained the basic information included in the ARLearn games. The presentations were delivered by the teachers during weekly classroom hours.

*Knowledge tests*. Background knowledge on energy domain was evaluated with a 17 fill-in-the-blanks and yes/no items from a K-8 school test on basic domain concepts as energy sources, fossil and renewable energy, energy efficiency of appliances, calculating efficiency. Changes in knowledge and understanding of issues related to energy consumption were traced with the help of a 5 items tests issued before and after the game.

*Motivation scale.* The intrinsic motivation (IMI) scale by Deci & Ryan [30] was used in the study. Three subscales were included: interest/enjoyment (5 items), effort/importance (5 items), value/usefulness (4 items) and perceived choice scale (3 items). Items were selected based on their relevance in the context from an open source instrument available at www.selfdeterminationtheory.org/questionnaires. As the items were repeated in the pre-test and post-test, they referred respectively to future or past activities. 5 items were formulated negatively and were reversed for analysis purposes. The items were translated into Dutch with minor textual and context-specific adjustments to the target group and the context of study.

An exploratory factor analysis was applied to understand the structure of the latent variable [31]. Internal consistency checks were done on the whole scale and the subscales.

*Pilot evaluation.* Both students and the participating teachers were de-briefed after the activity. Students filled in a three-item anonymous inventory on each of the 7 games they were offered through mobile devices. They indicated whether they had played each game, they rated each game they had played and could elaborate on the ratings. In a semi-structured interview the three participating teachers provided an evaluation of the game, discussed its contribution to school learning, and lessons learned. During the pilot the first author took notes of the classroom sessions and observations during activities in the school building.

### C. Design and Procedure

At the start of the pilot and one week after its completion paper and pencil questionnaires were administered (pre-post-test design). 75 students filled in both questionnaires. Both questionnaires included knowledge tests, and the motivation scale. After the pilot a short anonymous written student inventory and a group interview with the teachers were held.

*Game organization and progress.* Students started the game with a general introduction on the topic of energy consumption and an introduction of the tools (e.g. an app supporting data collection for their smartphones). Prior to the pilot a try-out was organized to let the students install an app on the personal mobile devices and get acquainted with some of the functionalities of this app. Informed consent request letters were distributed among the students for them and their parents.

The Energy game was organized as a 6-phase inquiry cycle, with each phase being introduced by a new mobile game. Two content specific games introduced relevant domain and methodological concepts. Thus, game One provided an introduction to the topic of energy consumption. Game Two introduced the inquiry cycle, the what's and how's of formulating research questions and hypotheses.

Led by the mobile games, students posed research questions in small teams, decided together how they would tackle the task, collected and analyzed data to answer the questions that they posed and presented the results to each other and the teachers in the form of posters during a poster walk. The poster walk was the object of a joint assessment activity: the teachers, present guests and researchers rated the presentations together with fellow students. Based on the ratings winners of the game were announced. The first, second and the third place teams were awarded 5-euro vouchers to an Internet game store.

*Teacher support.* Throughout the pilot the teachers and one researcher (first author) provided support at regular instructional moments in class. Scripted instruction through the ARLearn app guided data collection in the orientation phase (at home and at school) and in the data collection phase (at school) by means of audio, video, photo and text formats.

*Technical issues.* Throughout the pilot a number of technical problems have been experienced. The school WiFi-network which was necessary for the adequate functioning of the sensor network in the school building and for using mobile devices for personal inquiries proved less stable than expected. During several moments of the pilot, the

Wi-Fi-network was not available. The school technical support team directly related that to the sensor-network. Upon request of the school technical support team, a part of the sensor network was dismantled limiting the area in which energy consumption could be measured, influencing the number of inquiries that students could pursue.

# 4 Results

## 4.1 Effects on Student Motivation

As reported by Kalz and colleagues [25], analyses of variance of the complete scale scores indicated that the conducted activity did not succeed in keeping students interested – the total score on the motivation scale was significantly lower after the activity than during the first measurement. There was significant interaction effect of gender prior to activity. no difference in the motivation level could be seen in the post-test [25].

In order to better understand these results, analyses of the constituent components were attempted. To conduct comparative analyses at sub-scale level, sub-scales with a complete match at item level in both instruments had to be defined. First, a principal component analysis was done forcing a three factor extraction for both pre- and post-test scales and reliability tests on the emerged components were performed. Five items were eliminated from the scale in the process of reliability testing. The resulting sub-scales with a complete match at item level can be described as *interest/positive disposition* (6 items, Cronbach's α=.860 in the pre-test scale and .838 in the post-test scale, respectively); *perceptions of usefulness* (3 items, Cronbach's α=.765 and .838) and *effort/willingness to invest effort* (2 items, Cronbach's α=.745 and .861). The resulting 11-item complete scale was again subjected to principal component analysis which yielded satisfactory results.

**Table 1.** Self-reports on Motivation by Gender at the start of the activity

| | Interest (6 items) | Effort (2 items) | Usefulness (3 items) | Total motivation scale (11 items) |
|---|---|---|---|---|
| M | 18,8 (6,1)** | 8,2 (2,5)* | 13,1 (3,8) | 40,1 (9,6)** |
| F | 22,9 (7,7)** | 9,5 (2,9)* | 14,4 (3,9) | 46,7(12,4)** |

* $p< .05$, **p< .01;.

Table 1 gives an overview of the scale scores by gender in the pre-test and indicates results that were statistically significant. As could be expected, independent t-tests on the pre-test sub-scale scores point to significant differences between boys and girls as far as their general *disposition/interest* in the activity or *willingness to invest effort* in it are concerned. There are, however, no indications of gender differences in the perception of the task *usefulness*.

**Table 2.** Change in the motivation and constituent factors

| | Pre-test<br>*M (SD)* | Post-test<br>*M (SD)* |
|---|---|---|
| Interest (6 items) | 20 (3)* | 18,2 (7,2)* |
| Perceived usefulness (3 items) | 13,6 (3,9)** | 10,0 (4,3)** |
| Perceived effort (2 items) | 8,6 (2,6)*** | 7,6 (3,3)*** |
| Total scale (11 items) | 42,74 (10,8)**** | 36,38 (12,3)**** |

*$F(1,65)=6{,}922$, $p=.011$, $\eta p2=.096$; **$F(1,66)=43{,}784$, $p=.000$, $\eta p2=.399$;
***$F(1,66)=26{,}782$, $p=.000$, $\eta p2=.298$; ****$F(1,56)=5{,}712$; $p=.020$, $\eta p2=.093$,

Repeated measures analyses on the derived sub-scales point to loss of *interest, perception of usefulness of the task and perception of effort investment* for all participant. There is no effect on gender as a consequence of this activity. Boys and girls seem to be relatively unanimous in their attitude towards the activity. Table 2 provides an overview of the results and figure 3 illustrates the difference between the two patterns.

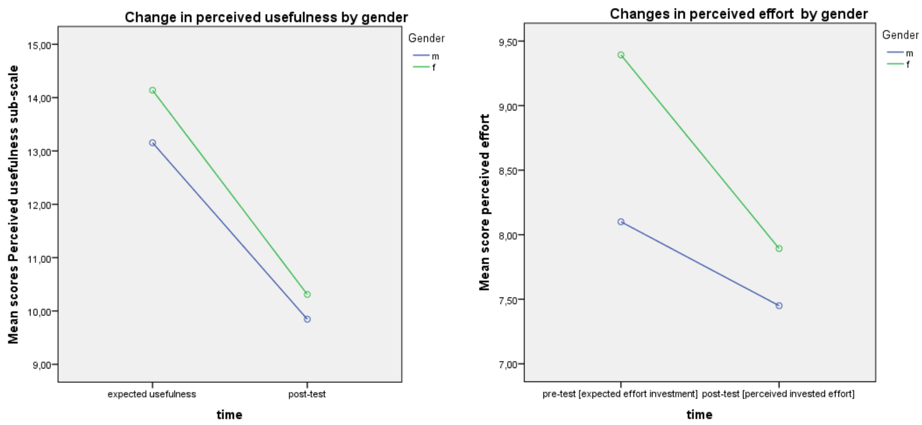



**Fig. 3.** Change in perceived usefulness and effort by gender

## 4.2 Pilot Process Evaluation

The idea of using mobile phones for learning was met with enthusiastic reactions (researcher's observation). As indicated in an anonymous evaluation inventory, practically all students used most of the ARLearn game scripts. According to their response (measured holistically with a single yes/no item) they demonstrated appreciation of activities that included actions (taking pictures, recording) and were negative on games that focus on supportive information. Understandably, students were most critical of inadequate functioning of WiFi network and the lack of structure because of last-minute changes in the activities that were meant to compensate for malfunctioning facilities.

In their interviews, the teachers acknowledged that they had accepted the suggested game scenario but did not contribute enough to its design and to structuring the activity. Last minute changes and ad hoc decision-making proved necessary to organize each game activity. As a positive moment, the teachers emphasized immediate feedback that students provided by demonstrating what they could or could not tackle independently after teacher introductions.

While there was a basic shared understanding of the concept of inquiry-based learning between the participants that have prepared the inquiry-activity (researchers and teachers), throughout the pilot it became clear that there was a wide range of understandings available with regard to the most important aspects of inquiry-based learning, i.e., the way students performed the inquiry activities and the teachers provided guidance.

## 5 Conclusions and Discussion

In the reported pilot study an inquiry-based approach, relating a topic within a school formal curriculum to students' immediate, familiar environments was realized in the form of a mobile serious game. Guided by direct prompts and indirect hints delivered through a mobile app, students pursued an inquiry around energy consumption, from defining a problem and a research question to the presentation of inquiry results.

The goal of the study was to evaluate the feasibility of the designed approach, to get insights in its practical constraints, and better understand the impact that the first implementation had on the learning outcomes, and in particular on learners' motivation. Answers to questionnaires on motivation point to a general decrease of interest, rather than to a boost of motivation, which is frequently reported in studies of mobile learning [23].

Furthermore, there are indications of gender differences as well as some subtle distinctions in the way students respond to questions related to "general interest" items, "willingness to invest effort" and "perceived usefulness". These results should be treated with caution as they are based on self-reports and the described factors are derived from a limited number of items. Nevertheless, they map into the theoretical perspective of the motivation and self-determination theory, namely the basic distinction between intrinsic and extrinsic motivation and the importance of the "inner acceptance of the value or utility of a task" [32,p.55] for actively engaging with the task and learning from it.

This point brings us to a relevant issue from the design perspective – the importance of letting students experience "usefulness" of tasks that they are performing. especially when students are introduced to a new pedagogical approach and new tools, like in this case when inquiry-based learning was combined with ubiquitous mobile devices.

**Acknowledgment.** The project was supported by the SURFnet innovation grant for sustainable ICT solutions. It made use of the technical solutions designed and developed in previous projects co-financed by SURFnet Innovation grants and the Welten

Institute – Research Centre for Learning, Teaching and Technology of the Open University of the Netherlands. It made use of an inquiry learning model and technological solutions developed in the weSPOT project. weSPOT is a 7th Framework Programme (FP7/2007-2013) project funded by the European Commission under grant agreement number: 318499 (http://wespot.net).

## References

1. Ally, M.: Mobile Learning: Transforming the Delivery of Education and Training. Athabasca University Press, Athabasca (2009)
2. Specht, M.: Learning in a Technology Enhanced World. OCÉ, Maastricht (2009)
3. Park, Y.: A Pedagogical Framework for Mobile Learning: Categorizing Educational Applications of Mobile Technologies into Four Types. International Review of Research in Open and Distance Learning 12(2), 78–96 (2011)
4. Campbell, S.W.: Perceptions of Mobile Phones in College Classrooms: Ringing, Cheating, and Classroom Policies. Communication Education 55(3), 280–294 (2006)
5. O'Bannon, B.W., Thomas, K.: Teacher Perceptions of Using Mobile Phones in the Classroom: Age Matters! Computers and Education 74, 515–525 (2014)
6. Ally, M.: Introduction. In: Ally, M. (ed.) Mobile Learning: Transforming the Delivery of Education and Training, pp. 1–6. Athabasca University Press, Athabasca (2009)
7. Sharples, M., Arnedillo-Sánchez, I., Milrad, M., Vavuola, G.: Mobile learning: Small Devices, Big Issues. In: Ludvigsen, S., Balacheff, N., de Jong, T., Lazonder, A., Barnes, S. (eds.) Technology-Enhanced Learning: Principles and Products, pp. 233–249. Springer, Dordrecht (2009)
8. Wong, L.-H., Looi, C.-K.: What Seams Do We Remove in Mobile Assisted Seamless Learning? A Critical Review of the Literature. Computers and Education 57(4), 2364–2381 (2011)
9. Ernst, J.: Factors Associated with K–12 Teachers' Use of Environment-Based Education. The Journal of Environmental Education 38(3), 5–31 (2007)
10. Zorgafakis, N., Menegaki, A.N., Tsagarakis, P.: Effective Education for Energy Efficiency. Energy Policy 36, 3226–3232 (2008)
11. Energy Literacy, Essential Principles and Fundamental Concepts for Energy Education US Dept. of Energy (2011), http://www.globalchange.gov
12. United Nations Environment Programme. GEO-s. Global Environment Outlook. Technical report, New York, NY (2012)
13. Volk, T.L., Cheak, M.J.: The Effects of an Environmental Education Program on Students, Parents, and Community. The Journal of Environmental Education 34(4), 12–25 (2003)
14. Ernst, J., Monroe, M.: The Effects of Environment-Based Education on Students' Critical Thinking Skills and Disposition toward Critical Thinking. Environmental Education Research 10(4), 507–522 (2004)
15. Rogers, Y., Conolly, K., Hazlewood, W., Tedesco, L.: Enhancing Learning: a Study of How Mobile Devices Can Facilitate Sensemaking. Personal Ubiquitous Computing 14, 111–124 (2010)
16. 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)

17. Kalz, M., Börner, D., Ternier, S., Specht, M.: Mindergie: A Pervasive Learning Game for Pro-Environmental Behaviour at the Workplace. In: Wong, L.H., Milrad, M., Specht, M. (eds.) Seamless Learning in the Age of Mobile Connectivity, Springer, Heidelberg (in press)
18. Van Rens, L., Pilot, A., van der Schee, J.: A Framework for Teaching Scientific Inquiry in Upper Secondary School Chemistry. Journal of Research in Science Teaching 47(7), 788–806 (2010)
19. Edelson, D., Gordin, D., Pea, R.: Addressing the Challenges of Inquiry-Based Learning through Technology and Curriculum Design. Journal of Learning Science 8, 391–450 (1999)
20. Littleton, K., Scanlon, E., Sharples, M.: Orchestrating Inquiry Learning. Routledge, London (2012)
21. Sandberg, J., Maris, M., de Geus, K.: Mobile English Learning: An Evidence-Based Study with Fifth Graders. Computers & Education 57, 1334–1347 (2011)
22. Song, Y.: "Bring Your Own Device (BYOD)" for Seamless Science Inquiry in a Primary School. Computers & Education 74, 50–60 (2014)
23. Schmitz, B., Klemke, R., Specht, M.: Effects of Mobile Gaming Patterns on Learning Outcomes: A Literature Review. International Journal of Technology Enhanced Learning 4(5-6), 345–358 (2012)
24. Ternier, S., Klemke, R., Kalz, M., van Ulzen, P., Specht, M.: ARLearn: Augmented Reality Meets Augmented Virtuality. Journal of Universal Computer Science 18(15), 2143–2164 (2012)
25. Kalz, M., Firssova, O., Boerner, D., Ternier, S., Prinsen, F., Rusman, E., Drachsler, H., Specht, M.: Mobile Inquiry-Based Learning for Sustainability Education in Secondary Schools: Effects on Knowledge and Motivation. In: Sampsons, D., Spector, M. (eds.) Proceedings of the The 14th IEEE International Conference on Advanced Learning Technologies - ICALT 2014: Advanced Technologies for Supporting Open Access to Formal and Informal Learning, Athens, Greece, July 7-9 (2014)
26. Protopsaltis, A., Seitlinger, P., Chaimala, F., Firssova, O., Hetzner, S., Kikis-Papadakis, K., Boytchev, P.: Working environment with social and personal open tools for inquiry based learning: Pedagogic and Diagnostic Frameworks. The International Journal of Science, Mathematics and Technology Learning (in press)
27. Tafoya, E., Sunal, D., Knecht, P.: Assessing Inquiry Potential: A Tool for Curriculum Decision Makers. School Science and Mathematics 80(1), 43–48 (1980)
28. Mikroyannidis, A., Okada, A., Scott, P., Rusman, E., Specht, M., Stefanov, K., ..., Chaimala, F.: weSPOT: A Personal and Social Approach to Inquiry-Based Learning. Journal of Universal Computer Science 19(14), 2093–2111 (2013)
29. Plugwise, `http://www.plugwise.com`
30. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior. Plenum Press, New York (1985)
31. Field, A.: Discovering Statistics Using SPSS, 4th edn. Sage, London (2009)
32. Ryan, R.M., Deci, E.L.: Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Contemporary Educational Psychology 25, 54–67 (2000)

# Technology Use in Lectures to Enhance Students' Attention

Vera Gehlen-Baum[1], Armin Weinberger[1], Alexander Pohl[2], and Francois Bry[2]

[1] Saarland University, Department of Educational Technology, Saarbrücken
v.gehlen-baum@edutech.uni-saarland.de
[2] Ludwig-Maximilians-University, Institute for Informatics, Munich

**Abstract.** Mobile devices, such as laptops, smartphones and tablets, are ubiquitous in lectures. Students report to use their mobile devices for lecture-related activities (e.g. taking notes). Observational data shows, that students use mobile device mainly for lecture-unrelated activities, like Facebook or playing games. So currently, mobile devices seem to distract learners from the lecture and ultimately hinder student-teacher interaction. In this study, we investigated how students (n = 75) use their mobile devices (N = 80) in a traditional lecture setting when supported with the technological support system "Backstage" or not. Backstage entails functions for quizzing students (Audience-Response-System) and a backchannel allowing students to interact with each other, commenting on slides, asking questions, and providing feedback to lecturers. The results show that this technology increases students' focus on lecture-related activities.

**Keywords:** lectures, related-activities, unrelated-activities, backchannel, ARS.

## 1 Mobile Devices in Lectures

Mobile devices are common in todays' lectures. While students claim to bring their mobile devices for lecture-related activities, like taking notes or searching for additional information, they often engage in lecture-unrelated activities, such as using Facebook, playing games or just surfing unrelated websites [1,2]. A negative correlation was found between lecture-unrelated mobile activities and academic performance [3,4,5]. Even though new forms of lectures are being discussed that partly integrate technology (e.g. flipped classrooms or inquiry learning), traditional lectures still prevail [2]. Traditional lecturing, with one lecturer talking most of the time, often fail to engage students in lecture-related activities and mobile devices seem to further distract students [2]. Lectures typically leave little time for including student activities [6]. Educational technology, like Audience-Response-Systems (ARS) or backchannels may support effective instructional practices in lectures, such as asking questions or providing feedback, which has shown to engage students in lecture-related activities [7,8]. In this study, we present different lecture-related and -unrelated activities students regularly engage in during lectures. In addition, we describe technological support, which can help to involve more students in the lecture. In an in vivo study,

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 125–137, 2014.
© Springer International Publishing Switzerland 2014

we observed students media activities when using mobile devices, with and without an additional learning platform, in a natural setting.

## 2 Students' Media Activities in Lectures

For students to build knowledge in lectures, they need to stay focused on the learning material [9]. When students bring their mobile devices to the lecture, they can use them for lecture-related activities, like taking notes or searching for additional material, which may help them to elaborate and understand the newly given information (see Tab 1). Students reported that their concentration continuously declines after the first 20 minutes until the end of the lecture [10]. Based on this finding, it is likely that students use their mobile devices also for lecture-unrelated activities. Lecture-unrelated activities, like surfing the internet and sending emails, distract students from the current lecture and forces students to switch their attention between different tasks [3], [11]. Students can focus their attention on their mobile device, for lecture-related and -unrelated activities. Students' attention is peripheral, when the mobile device is still switched on but the student is focusing on something other than the lecture, e.g., talking with a neighbor.

**Table 1.** Media activities in lectures [2]

| | Lecture-related | Lecture-unrelated |
|---|---|---|
| **Focused media activities** | - Note taking<br>- Searching<br>- Looking at slides | - Surfing the web (social networks, email, video etc.) |
| **Peripheral media activities** | - Background use of slides | - Background use of internet (social networks, email, video etc.) |

### 2.1 Lecture-Related Media Activities in Lectures

Lecture-related media activities can help students in lectures to understand the learning material better and therefore gain deeper understanding. Students can use their mobile devices to search for additional information or to take notes. To take notes, which fosters students' learning, students' attention has to be focused on the mobile device.

Some other, lecture-related media activities may not force the focus on the device, but demand peripheral attention only. One example is the lecturer offering handouts online, so students can open them on their mobile device. Instead of looking at those handouts on their laptops during the whole lecture, their attention is with the lecturer or in the communication with their neighbors.

### 2.2 Lecture-Unrelated Media Activities in Lectures

Besides the positive effects of bringing a mobile device for learning purposes, students can also use mobile devices for lecture-unrelated activities. Lecture-unrelated activities are those which distract students from the lecture content, e.g. social networks like Facebook, websites with unrelated content, emails or games. The results of a study of Kraushaar & Novak [4] with 97 students using spy-software indicates that especially the use of instant messages during lectures seems to correlate with lower learning results. One explanation for these findings is that students have to switch their attention between the lecture content and the lecture-unrelated activities, which reduces their cognitive resources [3], [11]. These cognitive resources should be further reduced when the lecture-unrelated activities are cognitively demanding and require attention. If the student focuses on the lecture-unrelated information in a peripheral way, for example by just opening Facebook while listing to the lecture, the switch between the different activities may be smaller and therefore the student will produce better (or less bad) learning results than when focusing on lecture-unrelated activities. Media usage in lectures may potentially indicate that students actively engage with lecture-unrelated content [3,4]. But so far, most of these studies use self-reports to investigate what students do with their mobile devices during lectures.

## 3 Additional Technology to Support Lecture-Related Student Activities in Lectures

Technologies in lectures can aim to guide students' attention as well as foster participation in lectures [7], [12,13]. Based on the assumption, that student engagement fosters motivation and learning results, backchannels and Audience-Response-Systems (ARS), have been developed in order to support students' learning in lectures and include more time for practice [8], [12], [14,15]. Making time for practicing what is being lectured seems to substantially improve learning results [16,17]. Unfortunately this kind of practice is rarely found in lectures [2], [18]. Technology may help to activate students' to participate in the lecture [7], [18]. So-called backchannels help students to interact with each other and the lecturer [8]. Audience-Response-Systems (ARS) offer the possibility to ask questions to a large number of students [12], [19]. While some of this technology has been developed specifically for lectures others had been adapted for this purpose.

### 3.1 Backchannel

Backchannels facilitate student activity in lectures as they help to establish public communication detached from the presenter's talk. The term backchannel describes additional communication during the lecture and refers also to the synchronous use of computer-mediated communication software. The lecturer typically dominates the frontchannel, presenting new information, and the number of student questions asked in a lecture is quite low (see [15], [20]). A technological supported backchannel can allow for anonymous communication, which may facilitate psychologically safe participation [21]. Building on social media technologies like microblogging or chats, different forms of backchannels can be integrated in a lecture. While some lecturers use Twitter to offer students the possibility to interact, other lecturers use platforms, which are designed especially for lectures [13], [22]. Platforms like Twitter solely support interaction without being integrated in the lecture. Alternatively, the backchannel is integrated in the lecture and students can interact with the lecturer by raising their hands virtually or post questions or annotations [23]. In some of those systems the backchannel, even though combined with other educational technology like ARS, is still a separated feature. It is also possible to connect students' questions and responses to the slides, so students can actively work with their learning material [24,25,26]. So far, there seems to be no study indicating that a specific form of backchannel leads to better learning results.

As especially students with suboptimal learning prerequisites tend to ask few questions in face-to-face settings [27], a backchannel may help to alleviate this problem. In a case based study students used a backchannel to ask questions which were already answered by the lecturer or more general questions [15]. The backchannel was a natural learning element of the university introduced by the students. Even though the backchannel was sometimes a distraction, it was accepted by the students and help them talking about content. When the backchannel is included by the lecturer, the results regarding engagement can vary. In a study with 23 students, Ebner and Schiefner [8] evaluated a microblogging tool with regard to its influence on learning. While microblogging accelerated communication, the number of people engaging in it was rather limited. As many backchannels are introduced by the lecturer it may be helpful to design backchannels carefully and to include further tools to facilitate student engagement [15], [28].

### 3.2 Audience-Response-Systems (ARS)

ARS were developed to support question asking of the lecturer. With an ARS the lecturer asks questions which can be answered by all the students in the lecture hall. Different student mobile devices, like smartphones, tablets, laptops or handheld devices handed to the students, offer all students the possibility to answer the lectures question simultaneously and anonymously. In a traditional lecture, the lecturer can ask a question for several purposes like assessing students' prior knowledge and their current understanding, or to foster elaboration on a certain topic [29,30,31], but only one student can answer the question [32]. This type of production blocking in lecture halls could be one explanation for the findings that higher order questions do not necessarily support better learning results [33], in spite of their beneficial effects on students' critical thinking and deeper elaboration [29,30]. ARS resolve the production blocking problem, since all students in a lecture hall are supposed to answer the question simultaneously [19], [34]. First, the lecturer asks a question, typically supported by a slide showing the question and the different answers in a written form. Second, all students can answer the question by pressing a hard- or software button, i.e. on a so called "Clicker" device or respectively, running a specific software on their mobile device [35,36]. The results of multiple-choice questions are often visualized as a bar-chart. These bar charts can be presented as feedback to all students simultaneously after the lecturer has stopped offering the question to the students. Questions in ARS can be used for the same purposes as traditional questions [29,30]. In addition, ARS can be used as a starting point for further discussions [12]. The results of a meta-analysis based on mainly self-reported data indicate that using ARS in lectures supports students' attention and participation [7].

## 4 Research Questions

RQ1: In which kind of media activities do students engage in during the lecture, with and without including lecture supporting technology?

RQ2: How does the flow of students' media activities change over the whole lecture, with and without including lecture supporting technology?

RQ3: Does the usage of lecture supporting technology in lectures improve students' lecture-related media activities?

## 5 Method

### 5.1 Participants

In this study, we observed 75 students with $N$ = 80 mobile devices (five students used two mobile devices) in three lectures (two of them randomly chosen) of the computer-sciences departments at the Ludwig-Maximilians-University (LMU), the Saarland University and the university of applied sciences HTW Saarbrücken. We analyzed 34 students in a traditional (29 male and five female) and 41 in a technology supported lecture (33 male and four female, for four students no data regarding gender was collected).

### 5.2 Learning Environment

Backstage is a learning environment developed in cooperation of the educational technology department of the Saarland University and the computer-science department of the LMU in Munich. The goal of Backstage is to foster students' learning activities by offering a possibility for more interaction as well as including lecturing instructions. The students' view of Backstage for this study can be divided into three segments (see Figure 1). Section A shows the backchannel activity, where students can see all messages and can select them with different parameters, like seeing only questions. Section B contains the slides offered by the lecturer. Students can interact with these slides by placing symbols, e.g., question or explanation-marks to indicate questions or comments. Here students also have the possibility to place their notes onto the slides. Section C offers a statement of the lecturer, like a summary or a comment for todays' lecture. For this study we decided to integrate two basic technologies in Backstage to create a foundation for later studies including scripts and learning analytics. Backstage supports the interaction of students with each other and the lecturer by including a backchannel. Therefore students can place a symbol on the slide and type their question in the backchannel on the left side next to the slide. Other students can follow the backchannel or click on a specific symbol to learn more of its content. Students also can answer the question or reply to a comment via backchannel. In addition, students are able to rate the quality of a question by pressing a "like" or "do not like" button. Also, the lecturer has the possibility to ask questions with an Audience-Response-System. Therefore the lecturer includes a multiple-choice question in Backstage between certain content slides. When the question appears in the slideshow, the lecturer starts the question and waits for a certain amount of time, while the students are answering the question. The students read the question on the slide in the middle of Backstage and choose one of the answers on the left side. After the lecturer stopped the question the results can be presented to the students. The results are often presented in form of a bar chart, where the right answer is pigmented in green and the wrong answers in red. Backstage presents for direct reading how many students answered the question and how many percent of them gave which answer.

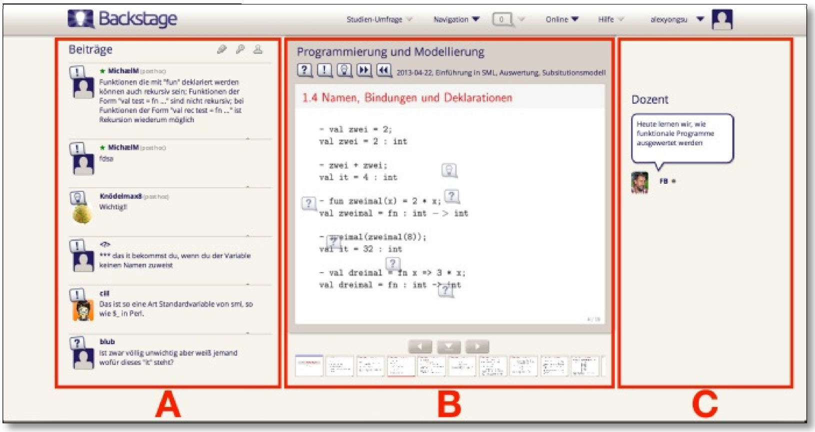



**Fig. 1.** View of backchannel in Backstage; Section A showing the backchannel, section B offers the lecturers slides to the students and section C contains a statement of the lecturer

### 5.3 Procedure

Before the beginning of the study two video cameras were placed in the lecture hall to video-tape the lecture activities. For the observations, five to nine observers were placed behind a group of one to four students with laptops in front of them. In both conditions, we used the same observers to ensure that the general procedure was similar. When the lecture started, the observers noted the on-screen activities every 30 seconds on a prepared sheet but were not allowed to take notes of any personal data.

At the beginning of the lecture, the lecturer or one of the researchers briefly informed the students about the study and pointed to further information at the end of the lecture. At the end of the lecture, the students got further information regarding the study and afterwards a questionnaire was handed to the students.

In the technology supported condition the lecturer and students had started using Backstage at the beginning of the semester and were already familiar with the system. The lecturer and the students received a short introduction in the beginning of the semester regarding the different functions of Backstage. Neither of them received an additional instruction how to use Backstage in order to gain better learning results.

### 5.4 Instruments and Variables

A prepared form, used for the observation, was categorized into lecture-related and lecture-unrelated mobile activities which had been observed in previous lectures. Lecture-related activity was coded when students took notes, searched for additional information, looked at lecture slides, or used Backstage. Lecture-unrelated activity was coded when students engaged in content-unrelated websites, emails, games or social networks like Facebook. Whenever the desktop or the starting page of a search engine was on the screen, neutral media activity was coded. Also displaying exercise sheets on the screen was coded neutral when it could not be observed whether these sheets concerned the current or other lectures, seminars, or homework. New activities, which had not been observed beforehand were added to the sheet and treated like the other activities.

Furthermore, we categorized media activities as focused or peripheral. When the students engaged in focused media activity, they interacted with their computers, like reading, writing or playing. Peripheral media activity was coded when the laptop was running, but the attention of the student was somewhere else, like writing something on a piece of paper or listening to the lecture. When the laptop was shut down, the coding was stopped or interrupted. The different coders were trained beforehand with accordance over 90 %.

In order to analyze how students participate in lecture-related or -unrelated activities during the lecture, it is necessary to reduce complexity of the data on the different activities like social networking or taking notes. All lecture-related focused media activity was therefore coded with 1 and lecture-unrelated activities with -1. Lecture-related peripheral media activity was coded with 0.5 and -unrelated behavior with -0.5. Neutral activities were coded with 0. This reduction of data allowed analyzing trajectories of students' media activities across a lecture.

The data were analyzed with regard to students' use of mobile devices using descriptive analysis. For differences in the conditions the Mann-Whitney-U-Test was used. For all statistical tests, the $\alpha$-level was set to 5%.

## 6 Results

RQ1: In which kind of media activities do students engage in during the lecture, with and without including lecture supporting technology?

In the traditional lectures, students spend nearly three times more time on lecture-unrelated activities 74.7% (*M* = 48.55; *SD* = 56.23) than they do on lecture-related ones 25.3% (*M* = 16.47; SD = 26.14). In addition, students' attention is more often focused on lecture-unrelated activities 59.8% (*M* = 38.89; SD =48.73), while students' attention on lecture-related activities is also often peripheral 10.5% (*M* = 6.84; SD = 16.64). For further division of students' attention in lecture-related and -unrelated activities see Table 1.

When integrating Backstage as a learning platform in the lecture, students spend 41.6% (*M* = 52.40; SD = 54.68) on lecture-unrelated and 58.4% (*M* = 73.48; SD = 68.96) on lecture-related activities. They focused on lecture-related activities 24.7% (*M* = 31.19; *SD* =39.67) of the time and 33.6% (*M* = 42.28; *SD* = 49.96) of the time were peripheral lecture-related activities (see Table 2).

In the traditional lecture six questions were asked by the students. In the technology supported lecture one question was asked on the frontchannel and two questions were asked during the lecture on Backstage.

**Table 2.** Frequency of exemplary lecture-related and unrelated media activities in lectures. (tech. supp = technology supported)

| | traditional - focused *M* (*SD*) % | traditional - peripheral *M* (*SD*) % | tech. supp. - focused *M* (*SD*) % | tech. supp. - peripheral *M* (*SD*) % |
|---|---|---|---|---|
| **Lecture-related** | | | | |
| Slides | 3.32 (8.62)<br>5.11% | 2.53 (8.59)<br>3.89% | 3.05 (8.23)<br>2.42% | 2.29 (8.19)<br>1.81% |
| Taking notes | 3.24 (8.10)<br>4.98% | 3.00 (14.44)<br>4.61% | 2.24 (7.16)<br>1.78% | .40 (1.25)<br>.32% |
| Backstage | - | - | 21.42 (34.49)<br>17.02% | 36.17 (49.62)<br>28.73% |
| **Lecture-unrelated** | | | | |
| Websites | 18.34 (27.35)<br>28.21% | 2.79 (6.00)<br>4.29% | 15.43 (16.96)<br>12.26% | 2.24 (5.55)<br>1.78% |
| Social networks | 8.13 (13.60)<br>12.50% | 1.55 (4.11)<br>2.38% | 9.60 (17.58)<br>7.62% | 1.26 (3.68)<br>1.00% |
| Games | 7.21 (32.03)<br>11.01% | .63 (2.54)<br>.97% | 8.81 (33.34)<br>7.00% | .57 (2.42)<br>.45% |
| Emails | .89 (1.71)<br>1.37% | .10 (.45)<br>.15% | 1.40 (3.06)<br>1.11% | .33 (1.59)<br>.26% |

RQ2: How does the flow of students' media activities change over the whole lecture, with and without including lecture supporting technology?

In the traditional lectures, students spend the majority of their time on lecture-unrelated activities. Only in a short period of time between minute 42 and 49, students engage more in lecture-related activities than in lecture-unrelated ones. At the same time, one of the lecturers included a break in his lecture. At the end of the lecture students activities tend to become even more lecture-unrelated (Fig. 2).

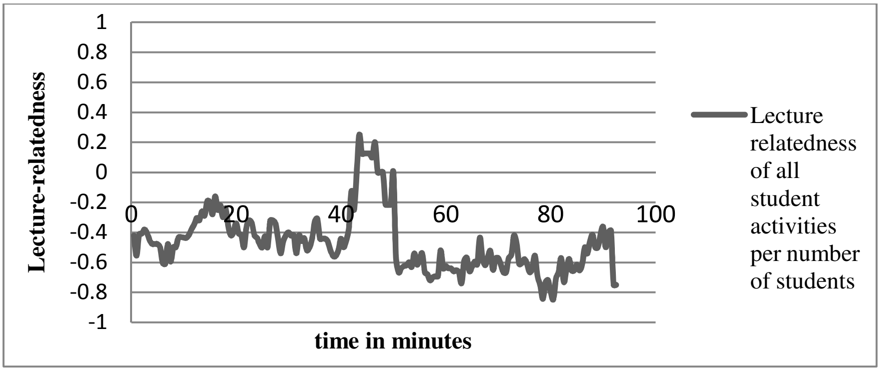



**Fig. 2.** Student activities during a traditional lecture

In the Backstage supported lecture, the students stayed on lecture-related activities for longer periods of time even though the lecture took 135 instead of 90 minutes. Also after the break, sometimes more students stayed on lecture-related activities than on -unrelated ones. Two questions were asked via the ARS by the lecturer. The first question was asked after 50 minutes and answered by 45 students. The second question was asked after 60 minutes and answered by 49 students. The posting of

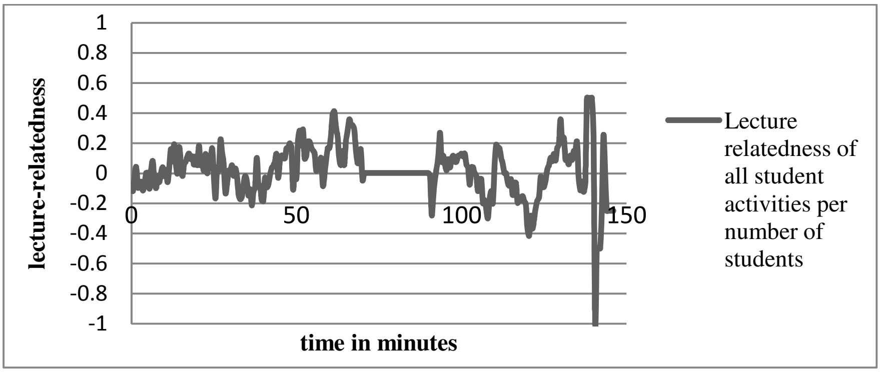



**Fig. 3.** Student activities in a Backstage supported lecture

questions are indicated by a small rise of lecture-related student activity (see Figure 3). Both questions were answered by approximately 2/3 of the students correctly. The fast decline, after students answer the Audience-Response-question, results of some students answering the question and checking lecture-unrelated content until the lecturer closes the question.

RQ3: Does the usage of lecture supporting technology in lectures improve students' lecture-related media activities?

While there is no significant difference in the amount of lecture-unrelated activities between students in traditional and technology supported lectures, $U = 743.00$, $p$ = n.s., students using Backstage spend more time on lecture-related activities $U = 405.00$, $p < .001$ (Fig.4). Similar results can be found for focused and peripheral student activities. Students in the technology supported condition spend more focused $U = 456.50$, $p = .001$ as well as peripheral time, $U = 418.50$, $p < .001$, on lecture-related aspects. The time they spend focused, $U = 749.50$, $p$ = n.s., and peripheral, $U = 792.50$, $p$ = n.s., on lecture-unrelated activities stays similar.

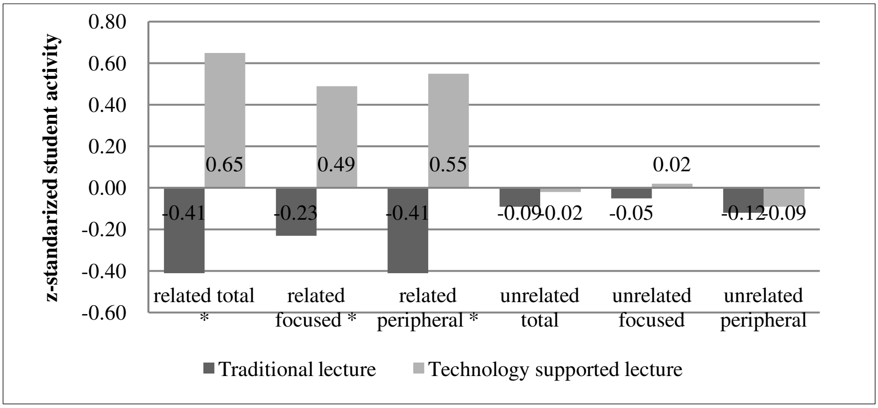



**Fig. 4.** Comparison of students' media activities in traditional and technology supported lectures (* = $p < .001$)

## 7 Discussion

The results of this study indicate that while students in the technology supported lecture seem not to ask more questions via backchannel, the overall lecture-related activity while using a mobile device increases. In a traditional lecture, there is just a very short period of time, where more students stay on lecture-related activities than on -unrelated ones. The students stay more time on lecture-unrelated websites, play computer games or use Facebook. In the backstage supported lecture, students used it most of the time. As the content of Backstage is mainly lecture-related, students focused and peripheral use of it may positively influence the learning results. In traditional lectures the concentration and lecture-related engagement decreases at the end of the session (see also [10]). Backstage seems to help students maintain their focus on lecture-related content for a longer period of time. Backstage does not change, however, that learners relate to the lecture using their mobile devices mostly in a peripheral way [13]. With or without Backstage, students tend to display slides, but do not interact with them, like taking notes on them. One possible explanation for the high peripheral focus on Backstage is that students wait and continuously check for some new peer comment to pop up, as has been found for email and Facebook use [5], [37]. Similar to other social media practices, only a few students may be actively producing content, like questions and comments, for the majority of students to consume passively on Backstage. Additional instructional support implemented into Backstage may raise the number of students producing content [30].

Even though including educational technology seems to have a positive influence on students' activity, just including a backchannel, where students can ask questions anonymously, seems not to lead to a higher number of questions asked by the students. While in the two traditional lectures six questions were asked, three questions were asked in the technology supported condition. Asking questions can be a useful tool to enhance students elaboration as well as higher thinking [29,30]. Based on the current results and the finding that students tend to ask factual instead of higher order questions [20], [30] additional instructions should be included [30]. Including educational technology offers the possibility to support students' questions when needed. One way to do so is to offer scripts or prompts to help students formulate higher order questions or tell students to ask questions anonymously [30], [38].

Addressing all students in a lecture with ARS may reduce the negative aspects of production blocking in lectures. In this study, approximately half the students joined in on answering the ARS questions. Further information with regard to student activities will be presented at the conference. At that time, students focused on the lecture, but also immediately used the idle time of waiting for all students to answer for lecture-unrelated activities, like checking emails. Ultimately, ARS may increase agency of students in lectures, may serve to orchestrate different learning activities, and advance the educational function of lectures to allow for participation and continuous assessment and feedback. Even though this study faces some methodical limitations, like the different duration of the lectures and the different lecturers, it demonstrated the trajectory of students' focus over a real-life lecture with or without supporting technology. Due to data privacy in this observational study, we could not collect personal data and relate our observations to learning results or students' perception of their media usage [1]. Further research should close that gap.

The results of this study indicate that including technological support in lectures activates students to spend more time on lecture-related activities and may therefore reduce the negative effects of mobile devices in lectures. In order to include educational technology productively in lectures it is therefore necessary to adjust the technology as well as the orchestration of the lecture. The technology can help for students to realize misconceptions in their knowledge and offer additional learning material to close those gaps. Therefore further research is necessary regarding the instructions of including students in the lecture with mobile devices and how that influences learning results.

## References


1. Gehlen-Baum, V., Weinberger, A.: Notebook or Facebook? How students actually use mobile devices in large lectures. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 103–112. Springer, Heidelberg (2012)
2. Gehlen-Baum, V., Weinberger, A.: Teaching, learning and media use in today's lectures. Computers in Human Behavior 37, 171–182 (2014)
3. Fried, C.B.: In-class laptop use and its effects on student learning. Computers & Education 50(3), 906–914 (2008)
4. Kraushaar, J.M., Novak, D.C.: Examining the affects of student multitasking with laptops during the lecture. Journal of Information Systems Education 21(2), 241–251 (2010)
5. Junco, R.: The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. Computers & Education 58(1), 162–171 (2012)
6. Faust, J., Paulson, D.: Active learning in the college classroom. Journal on Excellence in College Teaching 9(2), 3–24 (1998)
7. Kay, R.H., LeSage, A.: Examining the benefits and challenges of using audience response systems: A review of the literature. Computers & Education 53(3), 819–827 (2009)
8. Ebner, M., Schiefner, M.: Microblogging - more than fun? In: IADIS Mobile Learning Conference, pp. 155–159 (2008)
9. Renkl, A.: Aktives Lernen: Von sinnvollen und weniger sinnvollen theoretischen Perspektiven zu einem schillernden Konstrukt. Unterrichtswissenschaft 3, 197–212 (2011)
10. Stuart, J., Rutherford, R.: Medical student concentration during lectures. The Lancet, 514–516 (1978)
11. Kirschner, P.A., Karpinski, A.C.: Facebook® and academic performance. Computers in Human Behavior 26(6), 1237–1245 (2010)
12. Mazur, W.: Farewell, lecture? Science 323, 50–51 (2009)
13. Gao, F., Luo, T., Zhang, K.: Microblogging in education published in 2008 – 2011. British Journal of Educational Technology 43(5), 783–801 (2012)
14. Simon, B., Davis, K., Griswold, W.: Noteblogging: taking note taking public. ACM SIGCSE Bulletin, 417–421 (2008)
15. Yardi, S.: Whispers in the Classroom. Digital Media, 143–164 (2008)
16. Holt, L., Kysilka, M.: Instructional patterns. Strategies for maximizing student learning. Sage Publications, Ltd., Thousand Oaks (2006)
17. Hattie, J.: Visible learning. Routledge, Milton Park (2009)
18. Heaslip, G., Donovan, P., Cullen, J.G.: Student response systems and learner engagement in large classes. Active Learning in Higher Education 15(1), 11–24 (2013)
19. Campbell, J., Mayer, R.: Questioning as an instructional method: Does it affect learning from lectures? Applied Cognitive Psychology 759, 747–759 (2009)
20. Gehlen-Baum, V., Weinberger, A.: Question asking in large lectures. Presented at the EARLI 2013 (2013)
21. Fransen, J., Kirschner, P.A., Erkens, G.: Mediating team effectiveness in the context of collaborative learning: The importance of team and task awareness. Computers in Human Behavior 27(3), 1103–1113 (2011)

22. Junco, R., Heiberger, G., Loken, E.: The effect of Twitter on college student engagement and grades. Journal of Computer Assisted Learning 27(2), 119–132 (2011)
23. Wessels, A., Fries, S., Horz, H., Scheele, N., Effelsberg, W.: Interactive lectures: Effective teaching and learning in lectures using wireless networks. Computers in Human Behavior 23(5), 2524–2537 (2007)
24. Griswold, W., Simon, B.: Ubiquitous presenter: fast, scalable active learning for the whole classroom. ACM SIGCSE Bulletin (2006)
25. Hokanson, B., Hooper, S.: Integrating Technology in Classrooms. In: Anglin, J., Gary, E. (eds.) Instructional Technology. Past, Present, and Future, pp. 137–143 (2011)
26. Pohl, A., Gehlen-Baum, V., Bry, F.: Enhancing the digital backchannel Backstage on the basis of a formative user study. International Journal of Emerging Technologies in Learning 7(1), 33–41 (2012)
27. Brophy, J.E., Good, T.: Teacher behavior and student achievement. In: Wittrock, M. (ed.) Handbook of Research on Teaching. Collier Macmillan Publishers, London (1986)
28. Turban, G., Mühlhäuser, M.: A framework for the development of educational presentation systems and its application. In: Proceedings of the International Workshop on Educational Multimedia and Multimedia Education, pp. 115–118 (2007)
29. Barnett, J.E., Francis, A.L.: Using higher order thinking questions to foster critical thinking: a classroom study. Educational Psychology 32(2), 201–211 (2012)
30. King, A.: Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. American Educational Research Journal 31(2), 338–368 (1994)
31. Becker, G.E.: Unterricht planen. Beltz, Weinheim (2007)
32. Nassaji, H., Wells, G.: What' s the use of 'triadic dialogue'? An investigation of teacher - student interaction. Applied Linguistics 21(3), 376–406 (2000)
33. Winne, P.H.: Experiments relating teachers' use of higher cognitive questions to student achievement. Review of Educational Research 49(1), 13–50 (1979)
34. Brady, M., Seli, H., Rosenthal, J.: 'Clickers' and metacognition: A quasi-experimental comparative study about metacognitive self-regulation and use of electronic feedback devices. Computers & Education 65, 56–63 (2013)
35. Draper, S.W., Brown, M.I.: Increasing interactivity in lectures using an electronic voting system. Journal of Computer Assisted Learning 20(2), 81–94 (2004)
36. Gehlen-Baum, V., Pohl, A., Bry, F.: Assessing Backstage—A backchannel for collaborative learning in large classes. In: Proceedings of the Internation Conference ICL (2011)
37. Jackson, T., Dawson, R., Wilson, D.: The cost of email interruption. Journal of Systems and Information Technology 5(1), 81–92 (2001)
38. 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)

# Adding Epiphytic Assistance Systems in Learning Applications Using the SEPIA System

Blandine Ginon[1,2], Le Vinh Thai[1,3], Stéphanie Jean-Daubias [1,3], Marie Lefevre[1,3], and Pierre-Antoine Champin[1,3]

[1] Université de Lyon, CNRS
[2] INSA-Lyon, LIRIS, UMR5205, F-69621, France
[3] Université Lyon 1, LIRIS, UMR5205, F-69622, France
{name.surname}@liris.cnrs.fr

**Abstract.** In this paper, we present how the SEPIA system can be used to plug pedagogical and technical assistance systems in applications used by learners in an educational context. The SEPIA system consists in two main tools: an assistance editor that enables assistance designers to specify the assistance they wish for existing applications, and a generic assistance engine that executes the specified assistance in order to provide the application end-users with personalized assistance. We also present an experimentation of an assistance system setup with SEPIA in the context of a bachelor degree.

**Keywords:** User assistance, epiphytic approach, generic models.

## 1 Introduction

In the educational context, more and more applications are used, whether they are specifically dedicated to learning, the ILEs (Interactive Learning Environment), or other software used as material of a learning activity. The acquisition of knowledge by learners can nevertheless be compromised by difficulties of handling and use of the application. Adding an assistance system, suitable to both the application and the activity pedagogical goals, is a solution to face these difficulties and prevent the user from abandoning or under-exploiting the application and losing motivation [8].

The work presented in this paper is in the context of the AGATE project in which we have developed the SEPIA system (cf. Section 2.2). With the SEPIA system, assistance designers can add assistance systems to existing applications. This paper aims to present how the SEPIA system can be used to setup assistance systems for applications used in an educational context.

First, we present the AGATE project, its theoretical propositions and their implementation in the SEPIA system. Then, we present the actors and applications concerned by assistance in an educational context, before identifying the needs to which such assistance should answer. Finally, we explain how SEPIA can be used in the educational context and we detail the experimentation that we performed. We conclude by discussing the strengths and weaknesses of our propositions for assistance to learners, comparing them with other approaches.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 138–151, 2014.
© Springer International Publishing Switzerland 2014

# 2 Epi-assistance

This research work takes place in the context of the AGATE project (Approach for Genericity in Assistance To complEx tasks). It aims at proposing generic models and unified tools to make possible the setup of assistance systems in various existing applications, that we call *target-applications*. This project adopts a fully generic and epiphytic approach [14]. An epiphytic application, that we call *epi-application*, is an application able to perform actions in a third-party application without requiring any change to it. Thus, the functioning of an epiphytic assistance system doesn't disturb the functioning of the target-application. The models and tools proposed in the AGATE project are not specific to an application or to a domain, but on the contrary they can be used to setup assistance in a wide range of applications, without a need for these target-applications to have been designed specifically to enable the plug of assistance.

## 2.1 State of the Art of Epi-assistance Approaches

Several authors have studied the *a posteriori* specification of assistance systems for existing target-applications. The approaches of [15] and [7] make possible the plugging of an advisor system in a scenario from the Telos and ExploraGraph environments respectively. These advisor systems are defined by an assistance designer through a set of rules of the form <trigger event, trigger condition, assistance action, end event>. The trigger condition can include a consultation of the user profile and of the assistance history in order to contextualize the assistance. The proposed assistance actions are textual messages displayed in a pop-up for Telos, and animations or messages conveyed by an animated agent for Exploragraph. The approach proposed by [16] and by the CAMELEON model [4] make possible the plugging of an advisor system in a Web application, in order to trigger assistance actions when the end-user clicks on a link. The proposed actions are textual messages displayed in a pop-up, which can contain links to a Web page or to resources related to the assistance for [16], or an animated agent able to move, perform animations and display messages for [4].

Nevertheless, these different approaches cannot be used in any application. Indeed, they are specific to a given environment or to Web applications. In the context of the AGATE project, we are interested in the *a posteriori* plugging of assistance systems to very diverse existing target-applications. What's more, we would like to make possible a fine-grained personalization of the assistance, according to the user profile and to the assistance history, like in Telos and Exploragraph, but also according to the user's past actions, not only the browsing history as in [16], and according to the state of the target-application. Finally, to make possible a wider personalization of the assistance, we would like to propose a large choice of assistance actions.

### 2.2 Propositions: Process, Language and System to Add Epi-assistance

The AGATE project leads to several propositions that we present in this section: the adjunction process of epi-assistance systems in existing target-applications, the aL-DEAS language that enables the definition of various assistance systems and the SEPIA system that implements these theoretical propositions.

**Adjunction Process of Epi-Assistance Systems**

Figure 1 shows our adjunction process of epi-assistance systems organized in two phases: the assistance specification and the assistance execution in an epiphytic way.

The **assistance specification** is performed by an expert of the target-application, called the *assistance designer*. This preparatory phase enables the designer to specify the assistance that he wishes for a given target-application, expressed as a set of assistance rules. For this purpose, he must provide a description of the target-application interface in order to make possible the setup of an assistance that seems integrated in this application but executed in an epiphytic way. Then, the designer has to define the assistance rules that describe the behavior that the assistance system will have during the assistance execution.

The **assistance execution** concerns the target-application end-users. It consists in the execution of the assistance wished by the designer; it occurs at any use of the target-application by an end-user. This phase is composed of three processes: the monitoring of the target-application, the identification of an assistance need and the elaboration of an answer suitable to this need.

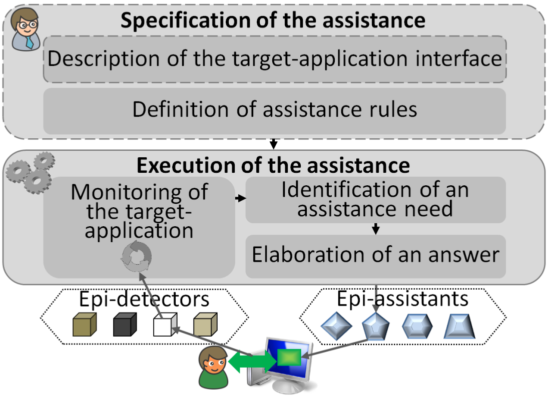



**Fig. 1.** Adjunction process of epi-assistance systems

The **monitoring of the target-application** is performed continuously while the user uses the target-application. It uses a set of epi-detectors (which are described in more detail in [9]) based on accessibility libraries that enable the access to all components of an application and the subscription to different kinds of events on these components. This process enables the epi-assistance system to get all information related to the interactions between the user, the target-application and the assistance system without disturbing the functioning of the target-application. The assistance system is aware of any user's low level action (like clicks, mouse movements, menu openings or item selections). This process uses the description of the target-application interface, built during the specification of the assistance, in order to identify the source component of each low level event. Indeed, knowing only the type (click, keystroke) of the user's action is not enough: it is much more informative for the assistance system to know precisely on which button the user clicked, over which image he moved the mouse pointer, etc., as in our proposition. While it is not the topic of this paper, let us stress the fact that the assistance designer could couple SEPIA to a trace management system [20] for combining low-level actions into higher level actions (*e.g.* filling a form or correcting red-eyes on a photo) that could in turn be used to trigger assistance rules.

The **identification of an assistance need** is performed when the monitoring process detects an event: it uses the assistance description made by the designer. More precisely, an assistance need is identified by a rule whose trigger event has just been detected by the monitoring process and whose trigger conditions are satisfied. In this case, the elaboration of an answer is launched.

The **elaboration of an answer** to the user's need uses the assistance description to trigger one or several assistance actions in the target-application with an epiphytic approach. This process also implements the personalization of the assistance according to the assistance designers' wishes. Indeed, the launched assistance actions can differ from a user to another according to his specificities, like his preferences for instance. The assistance actions are performed in the target-application by epi-assistants.

**The aLDEAS language**

aLDEAS (a Language to Define Epi-Assistance Systems) is a language that enables the definition of assistance systems as a set of rules. This language is the core of the adjunction process of epi-assistance systems described above: it enables the designer to define the wished assistance during the assistance specification phase, and it is used during the assistance execution through its implementation. aLDEAS is constituted of different types of components that can be combined to create assistance rules and actions, in order to answer the most various needs. The main elements of aLDEAS are event waits, consultations and assistance actions.

An **event wait** causes the execution to pause until the occurrence of the given event. An event can concern a user's action (like a click on a given button), an action of the assistance system (like the triggering of a given assistance rule), an absence of event (like an absence of click during a given duration), or the performing of a "high-level" task (like the red-eyes correction on a photo).

**Consultations** fetch information in order to personalize and contextualize the assistance. Thus, the triggering of an assistance action can depend on information relative to the user's past actions (his traces), to his choices, to the state of the target-application (in order for instance to know the content of a text field, or which item is selected in a combo box) and to his profile. The user profiles used in the AGATE project can contain any information found relevant by the assistance designer in order to personalize the assistance. It can contain in particular information on the user's preferences regarding assistance, or his skills in the target-application. For space reasons, how the user profile is constructed is not discussed in this paper.

Finally, aLDEAS provides a large choice of **assistance actions**: messages, component enhancements, automated actions on the target-application, proposition of external resources (like demonstration videos, forums or course materials), etc.

**The SEPIA System**

The SEPIA system implements these theoretical propositions through two main tools: an assistance editor and a generic assistance engine.

The **assistance editor** implements the assistance specification phase of the adjunction process of epi-assistance systems (cf. Figure 1). It provides assistance designers with a graphical interface for defining the assistance that they wish for a target-application, through a set of rules complying with the aLDEAS language.

The **generic assistance engine** implements the assistance execution phase of adjunction process of epi-assistance systems. It provides end-users with personalized assistance according to the rules defined by the designer.

The generic engine is completed by a set of **epi-detectors** that enables the target-application monitoring [9]. Currently, three epi-detectors have been developed: for Windows native applications, for Java applications and for Web applications.

Finally, the generic engine manages a set of **epi-assistants** that can perform assistance actions in the target-application at the request of the engine. The epi-assistants that we have developed can perform the assistance actions proposed by aLDEAS. In particular, they can display or read messages, enhance a component of the target-application interface, display an animated agent able to move on the screen, to express itself textually, orally, or by gestures and animations (for instance, the Merlin companion can applause and say "Well done, we succeeded!"). Some epi-assistants are also able to perform actions on the target-application interface, for instance to automatically click on a button, set text in a text field or select an item in a combo box.

aLDEAS also provides a set of *patterns* in order to facilitate the combination of aLDEAS elements, especially a pattern for assistance rules with the form <trigger event, optional condition, assistance actions, optional end-event>.

## 3 Assistance in an Educational Context

We wish to use the SEPIA system to setup epi-assistance systems in an educational context. More particularly, we would like to enable teachers to specify an assistance system for each target-application used by learners.

### 3.1 Which Actors Are Concerned?

In an educational context, the assistance designer is a pedagogical designer that wishes to plug an assistance system in an application that has no assistance or an assistance which is incomplete or unsuitable to his needs. The *pedagogical designer* can be a teacher or a pedagogical team, eventually assisted by a computer scientist or an expert of the target-application. The concerned end-users are the learners.

Another use of the SEPIA in the educational context could be the plugging of an assistance system by an assistance designer in a tool aimed at teachers as end-users, for example an authoring tool or a teaching assistant, but it is out of the scope of this paper.

### 3.2 Which Applications Are Concerned?

There are two main categories of applications used in an educational context: the ILEs and the other software used with a pedagogical goal. Both can be concerned by the adjunction of an epi-assistance system.

An **ILE** is an application that is specifically intended for learning. Otherwise, a "**non-pedagogical" application**, *i.e.* an application not specifically intended for learning, can also be used in the educational context. First, a pedagogical activity can aim at the discovery of an application. For instance, a company can organize a course to train its employees to use an application that they need in their work, like an Enterprise Resource Planning or Computer-Aided Design application. Some courses for the general public may also aim at teaching how to use a tool, like image editing applications or office applications.

Furthermore, learning how to use a "non-pedagogical" application is not the only pedagogical activity that may require the use of that application. A teacher can ask his students to use Word in order to learn how to present official letters in an office course, or to use Excel in order to learn algebra [17]. A tutor in a foreign language can also ask his course participants to use a sound recording application in order to improve their pronunciation. In these contexts, the non-mastery of the software used as course material can slow down the acquisition of the involved knowledge.

Finally, in some cases, the goals of a pedagogical activity can be about both application-related knowledge and domain knowledge. For instance, in a computer science course, a teacher can require his students to use a given IDE (Integrated Development Environment), such as NetBeans or Visual Studio, to perform exercises in order to learn how to use the environment *and* to acquire programming skills.

### 3.3 What Are the Assistance Needs?

The identification of the assistance needs is a key task for the specification of an assistance system. Thanks to a bibliographic study and to a study of existing assistance systems, we identified the main assistance needs of learners that use applications in an educational context (cf. Table 1). These needs are of two kinds: technical (T1 and T2) or pedagogical (P1 to P10). In this section, we illustrate these assistance needs with examples coming from research works.

**Table 1.** Assistance needs in learning

| Learners' assistance needs in an educational context | Description |
|---|---|
| T1. Handling | To help the first use of the application |
| T2. Use | To help the current use of the application |
| P1. Choice of the activity | To suggest activities suitable for the learner |
| P2. Learning prerequisites | To help the learner to acquire not mastered prerequisites |
| P3. Explanations on steps | To split up the different steps to perform for the activity |
| P4. Clues | To give clues helping to find the solution |
| P5. Examples | To give examples of similar problems or situations |
| P6. Transitional diagnostic | To add diagnosis as a complement to the application functionalities |
| P7. Explanations on errors | To explain his errors to the learner |
| P8. Sub-task automation | To perform a part of a task instead of the learner |
| P9. Summary, monitoring | To provide a monitoring during or at the end of the activity |
| P10.Pedagogical guidance | To propose a pedagogical scenario integrated in the application |

**Technical Assistance Needs**

Learners who use an application for the first time can face difficulties in handling (cf. T1-Table 1). They can also face difficulties in later uses of an application (T2), particularly in case of occasional use, or discovery of a new functionality. In general, these technical difficulties don't directly concern the knowledge to be taught and can also occur in a non-educational context. Nevertheless, they can prevent or slow down the acquisition of the involved knowledge or cause the learner to give up the pedagogical activity.

The setup of an assistance system to answer these technical assistance needs is one solution to face these difficulties. More particularly, in the educational context, such assistance systems prevent learners from losing time or motivation because of technical difficulties. What's more, in classrooms, a technical assistance system can make the learners more autonomous and thus enable the teacher to reduce his technical interventions in order to concentrate himself on the pedagogical aspects.

**Pedagogical Assistance Needs**

In addition to the technical assistance needs, learners that use an application in an educational context can face difficulties to which a pedagogical assistance could answer. The goal of a pedagogical assistance is not necessarily to enable a learner to finish correctly and quickly a pedagogical activity. To provide the solution to a learner seems not pertinent. However, some kinds of pedagogical assistance can be considered as pertinent by the teacher in order to facilitate the acquisition of the involved knowledge.

It is the case of the assistance for the choice of a pedagogical activity (P1), in particular when this choice is done by the pedagogical designer in a personalized way for each learner [11]. The help to acquire prerequisites (P2) can be useful to suggest to a learner to perform other activities (like exercises or courses) in order to help him in an activity for which he doesn't master all the prerequisites. For instance, the system

ELM-ART [3] presents to the learner links towards prerequisite concepts if he has never learnt or could not master them. Explanations on the steps to follow (P3) to perform an activity can be interesting to guide a learner and make him acquire methodology. For instance, in training mode, Aplusix [1] explains the steps of algebra problems solving upon request of the learner. Providing a learner with clues (P4, like in Aplusix) or examples (P5, like in Ambre-add [13]) can also facilitate the solving of the activity, the acquisition of the involved knowledge, as well as motivate the learner. The transitional diagnosis (P6) is useful to confirm to the learner that what he did is correct or to inform him that he did mistakes. Explanations on errors (P7) can be also given in order to help the learner to understand his mistakes. For instance, while Aplusix can only indicate that there is an error, Intelligent Tutoring Systems in Mobile Author [19] can give explanations on errors. In some cases, it can be relevant to automate a part of the task (P8) without compromising the acquisition of the involved knowledge. For instance, in the case of a difficulty in problem solving, Aplusix allows the sub-tasks automation upon request of the learner. A summary or a monitoring of the activity (P9) shows to the learner what he did and what he still has to do, in order to motivate him. For instance, the Explor@ Advisor Agent [12] displays the student's progression by indicating what has been achieved or completed. Finally, an assistance system can integrate a pedagogical scenario (P10) defined by the teacher and followed by the learner. For instance, OASIS [10] allows creating scenarios where learners are assisted step by step in the training mode, or where they try to achieve the final goal in a limited time without any help or feedback, in the evaluation mode.

As a conclusion, an assistance system used in an educational context may, if well designed, answer both technical and pedagogical assistance needs. The assistance provided to a learner must facilitate the acquisition of knowledge, complying with the designer's pedagogical strategy. In the next section, we present how the SEPIA system can be used to setup such assistance systems.

## 4 Use of SEPIA in an Educational Context

The SEPIA system has been designed mainly to enable the setup of assistance systems able to answer the technical assistance needs for the users of various target-applications. Nevertheless, we show in this section that SEPIA can also be used to setup assistance systems suitable to teachers' pedagogical strategies. First, we present examples of assistance systems specified with SEPIA for various applications used in educational contexts. Then, we detail the experimentation that we performed with SEPIA in the context of a bachelor course.

### 4.1 Use of SEPIA for Various Applications Used in an Educational Context

In French schools, 8 year-old pupils have to acquire skills in using a computer for the B2I diploma [2]. For this reason, primary teachers make them use different tools: some are specifically intended for learning, like the web site "Je Revise", or with a different purpose, like Paint. We have specified an assistance system for “Je Revise” that provides learners with pedagogical assistance and an assistance system for Paint in order to teach learners how to handle an image, which is one of the skills required for the B2I diploma.

Andes [18] is an Intelligent Tutoring System in physics problem solving. We have created an assistance system for Andes that guides students to choose activities in order to revise for an exam.

In the context of an ILE course in a master of University of Lyon, 20 assistance systems have been specified with SEPIA for a rudimentary ILE created by past students. These assistance systems aimed at helping 7-8 year-old children to discover the solar system using the ILE. Indeed this ILE didn’t provide any technical or pedagogical assistance.

We have created an assistance system for CodeBlocks that is used by students in the context of the first algorithmic course of University of Lyon. This assistance system guides students through all the steps of the first practical course.

We have also specified an assistance system to handle the forge (a platform allowing collaborative development) that is used by master students, and to teach them how to use SVN tools in order for instance to manage a project, to create a branch and to solve a conflict.

## 4.2 Experimentation of SEPIA for a HCI Course

We have also performed a deeper experimentation of the SEPIA system in an educational context: the setup of an assistance system in the framework of a bachelor degree course in HCI (Human-Computer Interaction), at University of Lyon. During this course, that is not a programming course; students have to design graphical applications, mainly in Java using the NetBeans IDE. However, many students have little experience with NetBeans and in Java programming and teachers notice each semester that students face difficulties that slow down their work, especially for the first course.

For this reason, teachers created demonstration videos available on the course website [5]. They also decided to define a tutorial integrated to NetBeans with SEPIA [6], in order to help volunteer students to discover the IDE and to practice basic exercises during the first course of the semester.

**Assistance Specification by Teachers**

In a first step, four assistance designers used the SEPIA assistance editor to specify an assistance system for the target-application NetBeans. This assistance system was a tutorial constituted of five independent parts corresponding to the main students’ assistance needs identified by the teachers during the first courses of past semesters: creating a NetBeans project, editing the properties of a NetBeans project, adding and using a button in a form, creating a menu bar and handling GUI events.

Each part of the tutorial contains around 14 assistance rules and 23 assistance actions, mainly of the type *assistance messages* with explanations or instructions, associated with an enhancement of the related components in the NetBeans interface. The creation of each part of the tutorial took about 3 hours. Although assistance specification represents a substantial work for teachers, the specified assistance system can be reused in future semesters.

As an example, Figure 2 shows a screen-shot of two assistance actions of the tutorial: a message and an enhancement of the text field "Arguments" in the screen "Project properties" of NetBeans. In this part of the tutorial, students learn how to retrieve and use these arguments in the parameter *args[]* of the *Main* function of the project. In this basic exercise, students have to use a *string array*, a *for* loop, and they have to cast a *string* to an *integer* and display a message in the *output*.

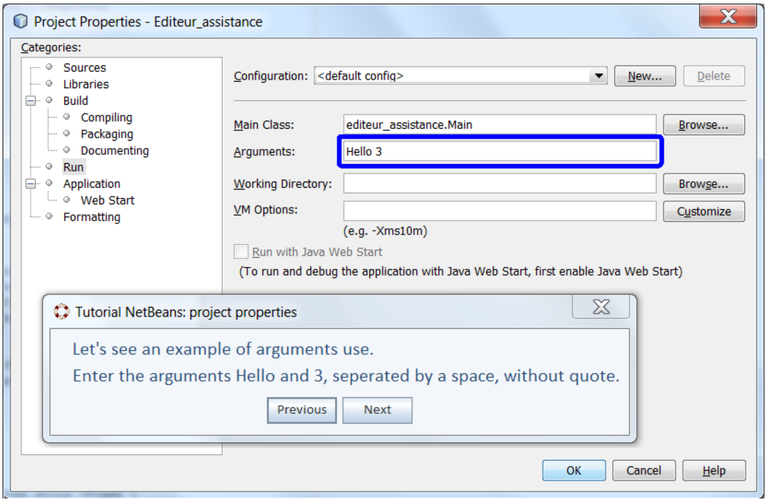



**Fig. 2.** Screen-shot of one SEPIA tutorial for NetBeans

**Assistance Execution for Students**

In a second time, teachers proposed to their volunteer students to use the tutorial at the beginning of the first course. On the 85 students, 64 volunteered to use the tutorial. Because of a lack of available computers with SEPIA assistance, only 52 students were able to participate. The tutorial sessions lasted between 30 and 90 minutes, depending on students. Indeed, students were free to work as they wanted with the tutorial: they could skip a part they deemed uninteresting and spend as much time as they needed on each part.

During the tutorial, students where free to exchange with other students and ask complementary information to teachers. Thus, we observed that some students who decided to skip a part of the tutorial changed their mind and do it after an exchange with another student who did it and found it useful. We also noticed that some students compared what they learned after each part of the tutorial. What's more, some of the students who had already used NetBeans compared with each other the method learned with the tutorial to the method that they previously used.

**Results of the Experiment with NetBeans**

The experiment was preceded by a questionnaire and a pre-test, in order to know the students' level in Java programing and their skill level in using NetBeans. After the definitive close of the assistance system, a final questionnaire and a post-test aimed at assessing their satisfaction and evaluating their progression concerning the target skills.

Students were very satisfied by the tutorial: 90.4% declared that they appreciated it (cf. left chart on Figure 3) and 62% wished to follow other similar tutorials integrated to NetBeans. In particular, in their comments on the final questionnaire, many students asked for tutorials integrated to NetBeans, but addressing more advanced Java programming skills, some students also asked for tutorials integrated to other tools and IDEs used in other courses. Moreover, most students found the tutorials more efficient than a demonstration video or a demonstration by the teacher (cf. Figure 4).

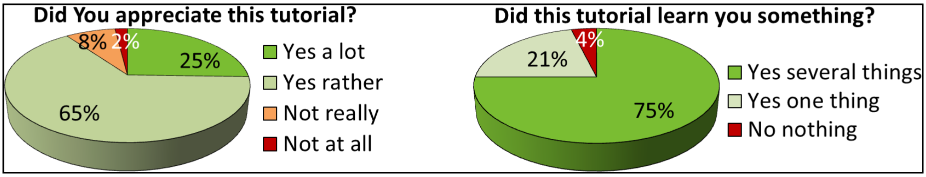



**Fig. 3.** Final questionnaire: students' satisfaction

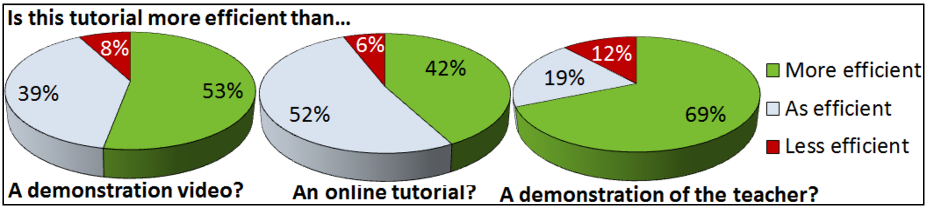



**Fig. 4.** Final questionnaire: students' opinion on the tutorial efficiency

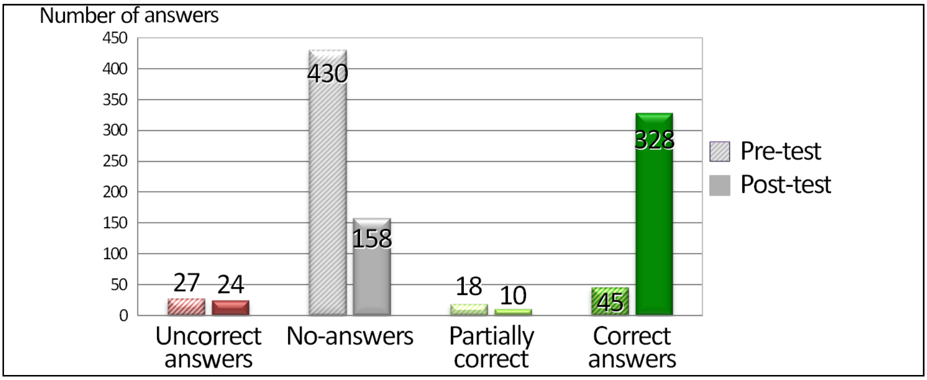



**Fig. 5.** Comparison of the results to the 10 questions common to the pre-test and the post-test, for the 52 participants

The tutorials seem to have facilitated the acquisition of knowledge and skills involved in the HCI course: 96.2% of the students declared that the tutorial taught them something (cf. right chart on Figure 3). This excellent result is confirmed by the improvement of their knowledge in NetBeans and Java programming measured by the comparison between the results of the pre-test and post-test (cf. Figure 5). Indeed, these tests contained notably 10 common questions with free answers: Figure 5 and Figure 6 concern these questions. We notice that the number of questions without answer has strongly decreased after the tutorial (- 63%), and that the number of good answer has very strongly increased (+ 629%). The post-test also contained 3 questions that wasn't in the pre-test, in order to see if students answered more correctly to the post-test only because they were more concentrate on the part of the tutorial that deals with the pre-test questions. For the post-test, the average success rate to the 10 common questions is 64%, and the average result to the 3 questions found only in the post-test is 62%, which is equivalent.

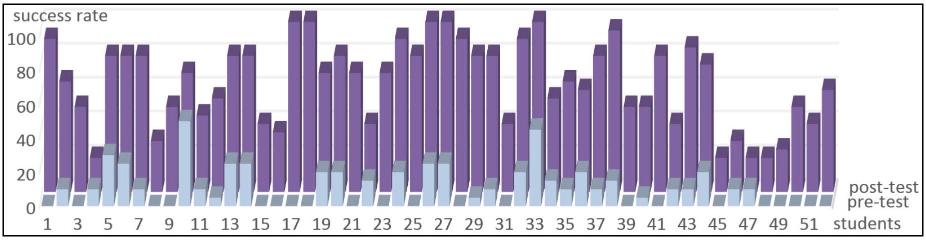



**Fig. 6.** Students' success rates to the 10 questions common to both tests

Figure 6 shows the success rate of each student to the pre-test (in the foreground) and to the post-test (in the background). We notice that all students improved their success rate after the tutorial: the smallest progression is 10%, the higher one is 100% (for students who had no correct answer to the pre-test and all correct answers to the post-test), and the average progression is 53.65%.

Furthermore, teachers were particularly satisfied by the use of the tutorial and plan to use it again for the next semesters. Teachers also would like to complete the tutorial with missing parts required by students in their comments. These parts, addressing more advanced Java programming skills, are not necessary for the first HCI course. However they can be useful for the rest of the semester, in the HCI course as well as for other courses that also require Java programming skills.

## 5 Discussion

This first experiment showed that the SEPIA system can actually be used to setup assistance systems in target-applications used by learners. Such assistance systems can meet the pedagogical goals of teachers and be well accepted by learners, like the tutorial for NetBeans used in a HCI course. Indeed, if the SEPIA system is particularly intended to setup technical assistance systems, it can also be used to setup pedagogical assistance system. However, there are some limitations to the use of SEPIA in an educational context that we discuss in this section.

First, for the moment, SEPIA cannot take domain knowledge into account. Such knowledge is yet necessary for most advanced assistance. For example, in a target-application that deals with algebra exercises, if the assistance designer wants to provide learners with an assistance that proposes transitional diagnosis, he has to specify all the possible correct answers (integrating for instance that "X-8" is equivalent to "-8+X"). It can be very demanding for the assistance designer and it becomes impossible in the case of an ILE that randomly generates the exercises. What's more, domain knowledge could be useful to generate examples and clues. Currently, the assistance designer specifies himself all the examples and clues necessary for his assistance system. With domain knowledge, the assistance system could automatically generate such examples and clues, contextualized with the current learner's production.

Then, knowledge relative to the activities proposed by an ILE would make possible an automation of some parts of the assistance designer's work. Indeed, for the moment, if an assistance designer wants the assistance system to guide learners through several activities, he has to specify the order of those activities. With the appropriate knowledge, the assistance designer would only have to specify that he wants the learners to perform the activities in the ascending order of difficulty. Moreover, with knowledge on the difficulty level of the ILE activities, the assistance designer could only specify that if a learner makes too many errors on an activity, then the assistance system should propose him for instance to move to an easier activity.

Finally, some improvements should be done to make the SEPIA assistance editor more accessible to assistance designers. Indeed, in the case of the HCI course, teachers were also computer scientists. However, in the general case, teachers may not be familiar with the concepts of rules, conditions, actions, and events that are at the core of the SEPIA system. Teachers could also face difficulties in understanding the description of the target-application interface that is an essential step to make a link between the epiphytic assistance system and the target-application. For the moment, SEPIA can be used by a teacher to setup an assistance system in an ILE, but with the help of a computer scientist, an expert of the ILE or a SEPIA expert.

## 6 Conclusion and Future Work

In this paper, we have presented the SEPIA system and its use to setup assistance systems in applications used by learners in an educational context. The assistance systems defined with SEPIA are epiphytic: they can be grafted on an existing application without a need to modify this application and without a need that the application has been designed specifically to enable the plugging of assistance.

We have shown how SEPIA can be used in an educational context to setup an assistance system suitable to teachers' pedagogical goals. This assistance system has been appreciated by learners and teachers that found it nice and useful. What's more this assistance system seems to have facilitated learning. We now plan to work on proposing patterns in order to simplify the work of assistance designers. In addition, the integration of domain knowledge to the SEPIA system would be an improvement to design assistance for ILEs involving complex problems solving.

## References


1. Aplusix, http://www.aplusix.com/en/
2. B2I diploma, http://eduscol.education.fr/cid46073/b2i.html
3. Brusilovsky, P., Schwarz, E., Weber, G.: ELM-ART: An intelligent tutoring system on World Wide Web. Intelligent Tutoring Systems, 261–269 (1996)
4. Carlier, F., Renault, V.: Educational webportals augmented by mobile devices with iFrimousse architecture. In: ICALT, Sousse, Tunisia (2010)
5. Demonstration videos for the HCI course, http://liris.cnrs.fr/~fduchate/ens/LIF14
6. Demonstration videos for the tutorial NetBeans, http://liris.cnrs.fr/blandine.ginon/PhDWork.html
7. Dufresne, A., Paquette, G.: ExploraGraph: A Flexible and Adaptive Interface to Support Distance Learning. In: Ed-Media, Victoria, Canada, pp. 304–309 (2000)
8. Gapenne, O., Lenay, C., Boulier, D.: Defining categories of the human/technology coupling: theroretical and methodological issues. In: Workshop ERCIM on User Interface for All, France, pp. 187–198 (2002)
9. Ginon, B., Champin, P.-A., Jean-Daubias, S.: Collecting traces in a Windows application. In: Workshop EXPPORT, ICCBR, New York, USA (2013)
10. Guéraud, V., Pernin, J.-P.: Developing pedagogical simulations: generic and specific authoring approaches. In: Lajoie, S.P., Vivet, M. (eds.) Artificial Intelligence in Education (AIED 1999). IOS Press (1999)
11. Lefevre, M., Jean-Daubias, S., Guin, N.: An approach for unified personalization of learning. In: UMAP Workshops (2012)
12. Lundgren-Cayrol, K., Paquette, G., Miara, A., Bergeron, F., Rivard, J., Rosca, I.: Explor@ Advisory Agent: Tracing the Student's Trail. WebNet, 802–808 (2001)
13. Nogry, S., Guin, N., Jean-Daubias, S.: AMBRE-add: An ITS to Teach Solving Arithmetic Word Problems. Technology, Instruction, Cognition & Learning 6 (2008)
14. Paquette, G., Pachet, F., Giroux, S., Girard, J.: Epitalk, a generic tool for the development of advisor systems. In: IJAIED, pp. 349–370 (1996)
15. Paquette, G., Rosca, I., Mihaila, S., Masmoudi, A.: TELOS: A Service-Oriented Framework to Support Learning and Knowledge Management (2007)
16. Richard, B., Tchounikine, P.: Enhancing the adaptivity of an existing Website with an epiphyte recommender system. New Review of Hypermedia and Multisedia, 31–52 (2004)
17. Ritter, S., Koedinger, K.: Towards lightweight tutoring agents. In: AIED, pp. 16–19 (1995)
18. Vanlehn, K., Lynch, C., Schulze, K., Shapiro, J., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M.: The Andes Physics Tutoring System: Five Years of Evaluations. In: AIED, Amsterdam, Netherland (2005)
19. Virvou, M., Alepis, E.: Mobile educational features in authoring tools for personalised tutoring. Computers & Education 44, 53–68 (2005)
20. Zarka, R., Champin, P.-A., Cordier, A., Egyed-zsigmond, E., Lamontagne, L., Mille, A.: TStore: A Trace-Base Management System using Finite-State Transducer Approach for Trace Transformation. In: MODELSWARD (2013)

# Individual Differences in Identifying Sources of Science Knowledge

Ilya Goldin[1], Maggie Renken[2], April Galyardt[3], and Ellen Litkowski[2]

[1] Center for Digital Data, Analytics, and Adaptive Learning, Pearson, Pittsburgh, PA, USA
ilya.goldin@pearson.com
[2] Educational Psychology and Special Education, Georgia State University, Atlanta, GA, USA
{mrenken,elitkowski}@gsu.edu
[3] Department of Educational Psychology, University of Georgia, Athens, GA, USA
galyardt@uga.edu

**Abstract.** We have developed an instrument to assess students' proficiencies in identifying sources of science knowledge (SoK) in text passages. We describe the new web-based instrument and our evaluation of the instrument with a sample ($n$ = 338) of children grades 2-8. By creating and validating this tool, we aim to establish a learning progression, inform science teaching, and tailor instruction to individual differences. Our findings suggest that students demonstrate differential ability in identifying SoK and thus imply the need for instruction to accommodate individual student perspectives on SoK. We expect that highlighting student ability in identifying SoK as a distinct skill will enable differentiated, adaptive instruction. We further expect this instrument to make explicit a component of what it means to think like a scientist, and in doing so facilitate conversations among teachers and students about the practice of science.

**Keywords:** individual differences, epistemic cognition, sources of knowledge, adaptive learning, formative assessment, science education.

## 1 Introduction

Overwhelming calls for improved student motivation and achievement in STEM education have led to emphasis on teaching students scientific practices in tandem with disciplinary content knowledge [1–3]. Additionally, recent reports such as the National Research Council's [4] *Learning Science through Computer Games and Simulations* emphasize the role of personalized approaches for improving students' science learning. Our research responds to these trends with a novel tool to measure a specific scientific process skill, to establish a learning progression, and to tailor instruction to students' individual differences. The current paper discusses first steps toward this tool. In what follows, we describe the development of a computer-based task measuring students' abilities identifying sources of science knowledge (SoK) and determine individual differences in their abilities to do so. Our consideration of individual differences takes into account situated, contextual factors such as disciplinary science content, particular source of science knowledge to be identified, and impact of training on the task. We validate the tool on cross-sectional data collected with grade 2-8 students to examine age-related trends.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 152–164, 2014.
© Springer International Publishing Switzerland 2014 2014

# 2 Related Work

## 2.1 Learner Epistemologies

The early development of learners' epistemologies, which encompass their understanding of science knowledge, may serve as an important foundation for later science learning [3]. We expect the ability to identify varied sources of knowledge is at the core of such sophisticated understanding. Conceptualization of learners' epistemologies, or more broadly epistemic cognition (EC), is scattered [5]; yet, a recent attempt to establish a guiding framework suggests

> "a more fine-grained analysis of students' epistemic cognitions—one more finely tuned to particular learning situations—can enable better explanation and prediction of learning processes and outcomes" [6].

The current work achieves a more fine-grained analysis in three regards. First, we narrow focus on a specific component of EC, namely *sources and justification of knowledge*. Referring to the sources of knowledge implies that knowledge originates in different manners and requires varied evaluation accordingly. For instance, knowledge can be passed down through the testimony of authority figures, or can be constructed via experimentation [6]. Developmental models reference a shift from knowledge as dictated by a higher authority to knowledge as self-generated [7].

Second, we consider ability across five science domains, and third, across five distinct sources of knowledge—observation, authority, cultural accumulation, theory testing, and reason.

## 2.2 Evidence Identification and Evaluation

Evidence evaluation has been examined extensively from two dominant research paradigms. Within the area of scientific reasoning, evidence is typically generated experimentally [8]. Alternatively, Harris and colleagues have widely considered young children's evaluation of testimonial evidence [9]. In everyday and in classroom contexts, students are not faced with either experimental evidence or the testimony of another person in isolation. Instead, students are charged with navigating, interpreting, and evaluating evidence from a variety of sources [10]. Developmental work suggests such navigation, interpretation, and evaluation are especially challenging tasks [11]. We propose that first, and in order to master the difficult task of evaluating evidence, students must be able to identify its source [12].

# 3 Methods

## 3.1 Data Collection

Participants ($n$ = 338) were from an elementary school (grades 2-5) and a middle school (grades 6-8) in the suburb of a large southeastern city in the United States.

54% of the participants reporting gender were female. The two participating schools are similar demographically, and ethnic/racial make-up is approximately 45% African American, 40% Caucasian, 5.5% Hispanic, 4.2% Interracial, and 2.5% Asian American, according to the school district website. Two classes per grade (2-8) were recruited for the current study. All classes at the elementary school and all science classes at the middle school were randomly assigned to participate in either the current study or simultaneous data collection for an unrelated study. Only students whose parents returned a participation opt-out form, or who were absent during data collection did not participate in the research study. As such, few students enrolled in the recruited classes were excluded from data collection. One student was omitted from the analysis, due to incomplete information.

Data were collected at one time point per class. Students were assessed on individual computers in a computer lab, in view of a trained research assistant. Before data collection, the research assistant walked participants through a practice task. During the practice, the research assistant read aloud and demonstrated how to select text in a passage. Participants could ask questions during this practice task.

### 3.2 Assessment Instrument

The instrument included five text passages. Each text explained a commonly accepted scientific phenomenon, starting with the question "How do we know…?" Topics included: "…plants grow," "…the Earth is round," "…objects fall when dropped," "…dinosaurs lived on Earth long ago," and "…water is a liquid." The texts were based on experts' explanations of the phenomena (collected in a separate study) and edited by the researchers. Experts ($n$ = 13) were graduate students or held a PhD in a relevant discipline.

This process allowed validity checking of the task in two ways. First, we compared SoK dimensions employed in expert explanations to our expected SoK dimensions, i.e., we checked that the sources of scientific knowledge that we considered in this study were also recognized as relevant to the phenomena by the experts, and vice versa. Second, we edited the texts to facilitate comprehension by the youngest participants. We piloted the task with a small cohort of second and fifth graders and asked students to think aloud during participation. Then, explanations were edited to be appropriate for grade 2-8 students, for example, by eliminating terminology that was misunderstood by students during pilot testing. Reading difficulty for each passage was estimated using the Reading Maturity Metric (RMM) [13]. Passages fell within Common Core Grade Bands as follows: K-1 for the plants passage, 2-3 for the earth and dinosaurs passage, and 2-5 for the water and objects passages. The word count averaged 129 words per text (min. = 89, max. = 200).

Texts were broken into segments. Each segment represented one of five SoK dimensions (observation, authority, accumulation, theory testing, and reasoning). For each text, participants responded to two kinds of items per SoK dimension: a "presence" item (one per passage per SoK dimension), and a "location" item (one per text segment per SoK dimension). The first "presence" item was a forced-choice, yes/no question. Participants indicated if the SoK dimension was present in the text. For example for the plants text and the observation SoK dimension, students read: "We may know that plants grow because we can see or observe it. Look at the paragraph. Did you read about seeing or observing plants grow?" Students who answered "yes" were directed to a second "location" item and asked to identify the specific segment of the passage that applies to observation. They read, "Click to highlight the sentences that are about seeing or observing plants grow." Notably, answering "no" to the presence item meant that the students were never asked any location items.

### 3.3 Design

Participants received the five science texts in a random order. SoK dimensions and sentences within each passage were presented in the same order across participants. Participants were randomly assigned to a training procedure or no-training. We considered that performance could be affected by familiarity with technology or comprehension of the user interface. In the training procedure participants were prompted to highlight a segment of the text aligned with an SoK dimension. For instance, they read: "We may know that objects fall when dropped because we can see or observe that objects fall. For example, the paragraph says, 'You probably see objects fall every day.' Click to highlight this sentence." The training procedure only included "location" items, as "presence" items were inappropriate for training. Participants in the training condition were randomly assigned to one of two training texts—the water or objects texts. Training passage was followed by four remaining texts in random order.

### 3.4 Modeling Epistemological and Age-Related Differences

We built several statistical models of the data. These models embody hypotheses about the learning progression and about individual differences. All model variants incorporate predictors representing student proficiency, student grade, passage difficulty, SoK difficulty, whether or not the student received a training procedure, and the unique id of the task that received the response. There are unique id's for each task, including each of the three different task types illustrated as items 1-3 in Fig. 1, and described in more detail below.

The data-generating process created by the current instrument meant that whether or not a student saw an item at the segment level (Fig. 1, no. 3) depended on their response at the paragraph level (Fig. 1, no. 2). In other words, if a student answered "no" to the question, "Is this SoK mentioned in this passage?," then they were not asked, "Which segment mentions this SoK?" Thus, there are structural reasons for why responses to some questions are missing, and we cannot treat these missing responses as missing at random.

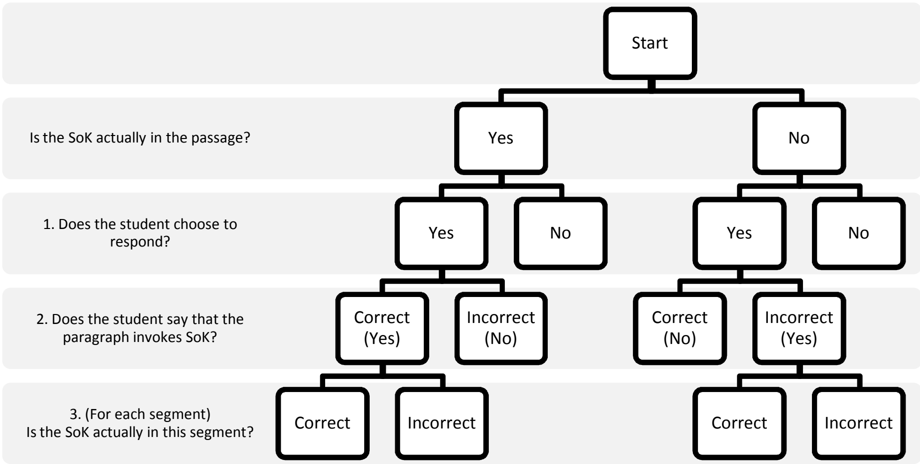



**Fig. 1.** IRTree decision tree

We address this issue with an IRTree modeling approach [14]. We can visualize the structure of the missing data with a decision tree (Fig. 1). Levels 1-3 in the tree represent different types of performance tasks; the models presented below let us investigate whether these tasks correspond to different dimensions of student ability. The response variable is a binary indicator of "success" on a given task. The three types of tasks are: success at choosing to respond at all, i.e., not skipping questions; success at recognizing whether or not a passage refers to an SoK; and success at identifying the SoK (or its absence) in each segment of the passage.

As IRTree is a Generalized Linear Mixed Model, this response is modeled as a logistic regression. We fit the models using a Bayesian Markov Chain Monte Carlo algorithm, as implemented in the MCMCglmm package in R [15]. The Bayesian formulation allowed us to "borrow strength" across students and tasks to obtain parameter estimates for factors with fewer data points, through shrinkage. We used uninformative priors for variables of primary interest, and weakly informative priors for student proficiency and item difficulty as described below.

The competing statistical models differ in how they represent interaction terms and student proficiency (Table 1). Models 1-3 are baseline models that include only the variables of primary interest. Models 4-9 test whether the relative difficulty of identifying the various SoK changes by grade level, by incorporating an interaction of grade level by SoK. Models 7-9 distinguish errors of omission and commission for the three different task types by including an interaction between an indicator of whether or not an SoK is present in the passage/segment and the type of task. That is, correctly *not* highlighting a segment that does not contain the SoK may be an easier task than correctly highlighting a segment that does contain the SoK.

**Table 1.** Parameters included in models 1-9. Proficiency is estimated using 1, 3, or 15 dimensions

| | proficiency-1 | proficiency-3 | proficiency-15 |
|---|---|---|---|
| **Baseline: grade, topic, SoK, training, item** | Model 1 | Model 2 | Model 3 |
| **Baseline + grade-SoK interaction** | Model 4 | Model 5 | Model 6 |
| **Baseline + SoK presence-task type interaction** | Model 7 | Model 8 | Model 9 |

The student proficiency terms took on one dimension (the "effect" of a student), 3 dimensions (one effect for each task type—respond, present, segment—for each student), or 15 dimensions (one for each task type, for each SoK, for each student). Proficiency was treated as drawn from a multivariate normal distribution, $\theta_d \sim MVN_d(0_d, \Sigma_{d\times d})$, where $d$ was 1, 3, or 15. The covariance matrix $\sum$ had an inverse-Wishart hyperprior, with an expected mean that is the identity matrix and $\nu = d + 3$. In the case of $d$ = 15, this meant we allowed the model to discover the levels of 15 kinds of performance from each student (one for each task type, for each SoK), as well as covariance among these levels. Further, the estimates of proficiency were partially pooled across the sample of students. That is, we "borrowed strength" from the sample average to shrink the estimates for those students for whom there were few observations, and we used the information on each individual student to compute the sample average, all at once.

Item difficulties were also modeled as coming from a common distribution, $N(0, \sigma^2)$, with a weak prior on variance $\sigma^2 \sim Inv.Wishart(V = 1, n = 1)$. The result is that the estimated difficulties of tasks with few responses, such as tasks in the training condition, are 'shrunk' towards the average item difficulty.

Each MCMC was run for 20,000 iterations, with a burn-in 3,000 iterations, and a thinning interval of 15 iterations. We verified that MCMC estimation of all of the models was stable by checking convergence of multiple chains, and autocorrelation for adequate mixing of chains.

# 4 Results

## 4.1 Epistemological and Age-Related Difference

To compare these non-nested models, we use the Deviance Information Criterion, DIC. Similar to AIC and BIC, but tailored to Bayesian modeling with MCMC estimation, DIC rewards models for accurately predicting the response, but penalizes models for using too many parameters, i.e., for overfitting the data.

First, models 9, 3, and 6 outperformed all other models, implying that the 15-dimensional representation of student proficiency contributed to predictive accuracy better than representations with fewer dimensions. Models 8, 5, and 2 with the 3-dimension representation of student proficiency for different task types strongly outperformed the 1-dimensional representation of proficiency.

**Table 2.** DIC scores for all 9 models (lower is better). * indicates the best model.

| | proficiency-1 | proficiency-3 | proficiency-15 |
|---|---|---|---|
| Baseline: grade, topic, SoK, training, item | Model 1<br>58427.24 | Model 2<br>53828.20 | Model 3<br>53240.57 |
| Baseline + grade-SoK interaction | Model 4<br>58421.17 | Model 5<br>53820.39 | Model 6<br>53259.06 |
| Baseline + SoK presence-task type interaction | Model 7<br>58397.57 | Model 8<br>53769.44 | Model 9*<br>53212.16 |

Second, in reporting the results for the baseline variables, we focus on the best fitting model 9. However, it is important to note that results for the baseline variables were similar in all models. There were no differences in performance by grade, that is, the 95% credible intervals (CIs) overlap substantially for all of the grades. Among the SoK dimensions, observation was significantly harder than accumulation. The difficulty of other SoK dimensions, authority, reason, and theory testing, was not discernibly different than observation or accumulation. The passage topics did not differ in difficulty. The training procedure made the training topics significantly easier, but the effect of training did not persist to other topics.

Third, the interaction of grade by SoK did not contribute consistently to predictive accuracy: models 5 and 4 (including the interaction), outperformed models 2 and 1 (omitting the interaction), respectively, but model 6 performed worse than model 3. In fact, model 3 was the second-best performing model.

Fourth, the interaction of true SoK presence by task type contributed consistently to predictive accuracy; model 9 outperformed models 3 and 6, model 8 outperformed models 2 and 5, and similarly for models 7, 4, and 1. Examining the parameters more closely, at the "respond" (highest) level of the IRTree, whether or not an SoK was present in the passage did not affect the response. At the second level of the IRTree, passages where an SoK was actually absent were less likely to elicit a correct response than passages where an SoK was actually present. This may be because students inferred that each passage actually invoked most or even all SoK. Finally, the level of passage segments was significantly more difficult than the level of entire passages. Items on *segments* where an SoK was actually absent were about as difficult as items on *passages* where an SoK was actually absent; however, items on segments where an SoK was actually present were much more difficult.

Finally, we examine the covariance of the 15-dimensional student proficiency matrix (Fig. 2). There is a very strong correlation between performance with all SoK at the top ("respond") level of the IRTree, i.e., choosing to respond (or not) does not depend on the SoK. Correctness at the "present" passage level is also highly correlated across the SoK. Correlation across the SoK at the segment level is lower, but still substantial. Though, we note that observation was the hardest SoK and proficiency in observation at the segment level has a low correlation with other proficiencies.

Proficiencies did not differ between grades at the passage and segment tasks. However, at the skip/respond task, older students had much higher proficiency estimates (Fig. 3). In other words, younger students, including the 6$^{th}$ graders, were much more likely to skip an item entirely than older students, but the younger students who *did* respond performed at the same level as older students.

Proficiency for the segment tasks has a moderate negative correlation with proficiency at the passage-level tasks. This may be due in part to the fact that "Yes" was the correct answer for 21 of the 25 passage-level questions, so a student could guess quite easily. Guessing is a much less effective strategy on the segment-level tasks.

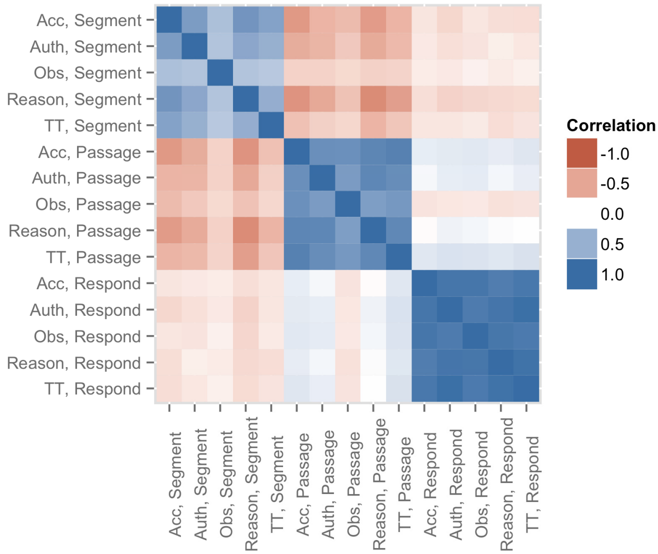



**Fig. 2.** Student proficiency correlation matrix from interaction of SoK by IRTree level

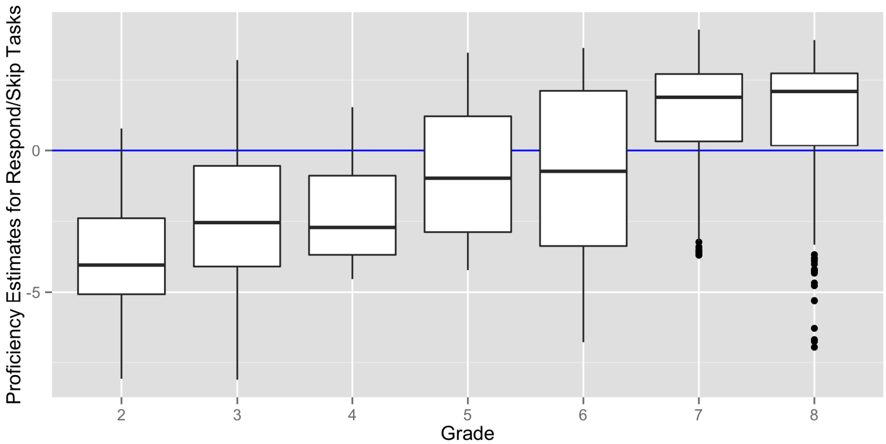



**Fig. 3.** Student proficiency estimates for the respond task across grade. A proficiency estimate of zero indicates the average probability to respond across all 337 students.

### 4.2 Post Hoc Analysis of Easy and Difficult Text Segments

We inspected the estimated item difficulty parameters to identify any additional factors that may influence the difficulty of the task. The passage on the earth topic (Fig. 4) is illustrated here because this passage contains some of the easiest and most difficult items after controlling for all other variables, including whether incorrect answer involves an error of omission, i.e., not selecting the correct SoK, or commission, i.e., saying it is present in a block of text when it is not.

Many of the easiest items are the first and last sentences in a passage, including the first segment of the earth passage. The harder segments are located primarily in the middle of the passage, appear to be longer, and include both errors of omission and commission. This seems to indicate that difficulty in identifying whether an SoK is present in a segment may be related to the reading difficulty of the segment, as well as how easy the segment is to visually locate within a passage.

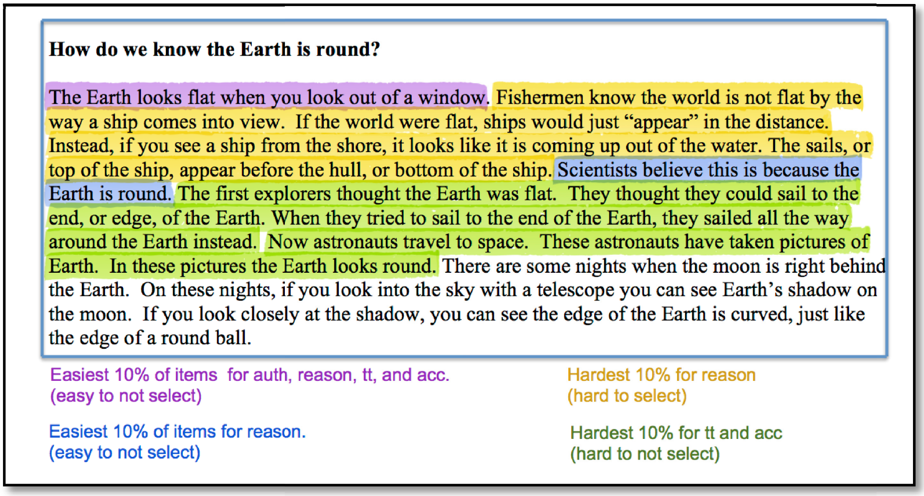



**Fig. 4.** Easy and difficult segments of text in the earth passage (best viewed in color)

## 5 Discussion

The current work provides a first step in creating a tool to measure the specific scientific practice of identifying multiple, valid sources of scientific knowledge. This study is part of a larger research program aiming to create and validate such a measure, to establish a learning progression, and to tailor instruction to students' individual differences. Our findings assist in achieving these aims, highlighting task difficulties and individual differences. We discuss the contribution of these findings to the broader literature base related to learners' abilities identifying SoK and to the future development and validation of the individualized web-based tool.

### 5.1 Identifying SoK as a Domain-General Ability

Our rationale for creating a measure of SoK identification ability is grounded in a framework of epistemic cognition. Increasing criticism of measures of epistemic cognition centers on traditional disregard for the situated nature of learner epistemologies [5]. One such situated context may be the domain of knowledge in question, e.g., physics, biology, or chemistry. Traditional approaches target domain-general epistemologies; however there is evidence that individuals' beliefs related to knowledge are domain-specific [16, 17]. We considered the possibility of domain-specific abilities in identifying SoK by including topical passages based in five different science domains in the task. Student ability did not differ from passage to passage, suggesting that identifying SoK in this task is a domain-general ability with regard to science discipline.

### 5.2 Difficulties Identifying the Observation Dimension

The instrument allows consideration of five distinct SoK dimensions. Unlike the uniform student ability across passage topics, SoK dimensions themselves are not uniformly easy to identify. Overall, SoK dimension constructs are stable across passages, but identifying observation is more difficult than identifying cultural accumulation of knowledge. Several possible interpretations of observation's difficult status exist.

First, this finding may be a side effect of an inadequacy in the instrument itself. The segments of text prescribed as observation may have been more difficult at a text comprehension level. Second, students may be less familiar with observation as an SoK than they are with other sources of knowledge. This is unlikely, because students engage in acts of observing across domains in their everyday lives. Third, the SoK location items were administered in a fixed order. Observation could be more difficult because it was always presented as the first SoK in a passage. If a student believes that only one SoK applies to a given sentence, then the final SoK in the SoK sequence, in this case reasoning, may be "easier" because segments have been eliminated from a potential bank of answers. This interpretation also is unlikely; difficulty associated with the last presented SoK dimension, reasoning, was no different than other dimensions.

The final, and most likely, interpretation is that the observation as SoK construct may be the most poorly defined. Observation is required in testing a theory and is often the basis for an authority figure's testimony. Such overlaps in observation and other dimensions, e.g., theory testing and authority, combined with classroom emphasis on experimentation and the testimony of more knowledgeable others, e.g., a teacher, may lead students to not conceptualize observation as a distinct SoK dimension. We will further examine these interpretations in future work

### 5.3 Early Preparedness for Identifying SoK

Student ability identifying SoK is stable across grade level within SoK dimensions. The lack of an age-related trend may seem unexpected. With the typical development of greater cognitive ability with age, we may expect performance on the current task to improve in tow. However, it is more likely the case that while *evaluating* SoK relies on higher order cognitive abilities, *identifying* SoK is a fundamental ability with potential to come on line early in childhood. As demonstrated here, students are developmentally prepared to identify varied SoK in text as early as second grade.

### 5.4 Individual Differences in Responding, Deciding Presence, and Locating SoK

Our findings illuminate individual differences not best represented in an overall ability to identify SoK, but rather in responding, deciding if an SoK is present, and identifying a prescribed SoK in segments of text. In particular, there is a clear tendency to avoid responding in the lower grades. Paired with no difference in ability locating SoK in segments of text across grade, this may imply on the one hand that older students may have misplaced confidence in their ability, or, conversely, a propensity to provide answers even if they are guesses. On the other hand, it may be the case that younger students have either some metacognitive understanding of their limited expertise or inappropriately low confidence in their abilities.

The presence of this “skipping” effect, especially among younger students, and differences between the position of the hard and easy items necessitates some refinement of the task. For instance, we may randomize the order in which text segments are presented in the passages themselves.

### 5.5 Limitations and Future Directions

The current study did not include a separate reading measure to account for individual capability with text comprehension. Although reading comprehension may account for some of the demonstrated individual differences, we edited the passages for readability based on pilot data, and RMM scores suggest the passages were reasonable for use with grade 2-8 students. The current study did not include a separate measure to account for individual proficiency with technology. However, it is unlikely that proficiency with technology impacted student ability, because strong student performance on training passages suggests that students were able to manipulate the user interface. Because the training procedure led to strong performance on training passages but not subsequent passages, this training paradigm should be improved.

Future work will address revision of the instrument through both additional modeling and experimental methods. Although the current study does not suggest an age-related trend, consideration of a developmental trend within the individual is not included here. Future models should explore developmental, or learning, trajectories, including passage order as a predictor of student ability. Other covariates for inclusion in future models include student reading comprehension. Because the reading maturity differed across texts, fitting a model with a proficiency-by-topic interaction could also serve as a proxy for text comprehension. Future studies of modifications to the instrument will consider randomizing SoK dimensions within passages, presenting SoK segments as bulleted lists rather than in paragraph form, and the impact of adaptive tutoring in place of the current training paradigm for improved performance.

## 6 Conclusions

A primary goal of the current work is to consider learning trajectories related to elementary and middle school students' understanding of sources of science knowledge. In exploring this broad understanding, the larger research program considers learners' abilities in identifying multiple, valid sources of knowledge and their evaluations of these sources. As a first step toward testing this assumption, the current study narrowed focus to individual differences in ability to identify prescribed sources of science knowledge across several topical domains. Our findings suggest identifying sources is a skill present as early as second grade and does not improve markedly through middle school. Our findings also demonstrate variation in student ability within grade levels, suggesting the need for personalized approaches to develop this fundamental skill, especially in older students. This research is valuable for informing such approaches as well as informing our understanding of learner epistemologies related to sources of knowledge.

**Acknowledgments.** The work presented here was funded in part by the University System of Georgia Board of Regents Arts and Sciences STEM Initiative.

## References

1. Forsthuber, B., Motiejunaite, A., de Coutinho, A., Sofia, A.: Science Education in Europe: National Policies, Practices and Research. Education, Audiovisual and Culture Executive Agency, European Commission (2011)
2. Council of the European Union: Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training. Official Journal of the European Union C. 119, 2–10 (2009)
3. National Research Council (U.S.): A framework for K-12 science education: practices, crosscutting concepts, and core ideas. The National Academies Press, Washington, DC (2012)
4. Honey, M.A., Hilton, M.: Learning Science through Computer Games and Simulations. National Academies Press (2011)
5. Renken, M.D., Chinn, C.A., Vargas, P., Sandoval, W.: Exposing and Assessing Epistemic Thinking. In: Proceedings of 11th International Conference of the Learning Sciences, Boulder, CO (2014)
6. Chinn, C.A., Buckland, L.A., Samarapungavan, A.: Expanding the Dimensions of Epistemic Cognition: Arguments From Philosophy and Psychology. Educational Psychologist 46, 141–167 (2011)
7. Hofer, B.K., Pintrich, P.R.: The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. Review of Educational Research 67, 88–140 (1997)
8. Zimmerman, C.: The development of scientific thinking skills in elementary and middle school. Developmental Review 27, 172–223 (2007)
9. Harris, P.L.: Trusting What You're Told: How Children Learn from Others. Harvard University Press (2012)

10. Next Generation Science Standards: For States, By States. The National Academies Press (2013)
11. Kuhn, D., Cheney, R., Weinstock, M.: The development of epistemological understanding. Cognitive Development 15, 309–328 (2000)
12. Renken, M.D., Litkowski, E., Oliver, M., Goldin, I.M.: Students' recognition of the value of scientific process when evaluating science explanations. Poster at American Psychological Association Annual Conference, Washington, DC (2014)
13. Landauer, T.K., Kireyev, K., Panaccione, C.: Word Maturity: A New Metric for Word Knowledge. Scientific Studies of Reading 15, 92–108 (2011)
14. De Boeck, P., Partchev, I.: IRTrees: Tree-Based Item Response Models of the GLMM Family. Journal of Statistical Software 48, 1–28 (2012)
15. Hadfield, J.D.: MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software 33, 1–22 (2010)
16. Estes, D., Chandler, M., Horvath, K.J., Backus, D.W.: American and British college students' epistemological beliefs about research on psychological and biological development. Journal of Applied Developmental Psychology 23, 625–642 (2003)
17. Renken, M.D., McMahan, E., Nitkova, M.: Initial validation of an instrument measuring psychology-specific epistemological beliefs. Teaching of Psychology (in press)

# Digital Didactical Designs of Learning Expeditions

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}@umu.se

**Abstract.** Current studies on media tablets illustrated that mobile technology may improve learning when truly integrated into learning settings. The question remains what truly integrated means, how it might be operationalized. In a study of Scandinavian classrooms, the question was how teachers adopt and integrate media tablets in their teaching practices in order to provide learning opportunities for their students. Seven K-9 schools implemented media tablets for around 2,000 students aged 6-16 and ca. 170 teachers in a 1:1 programme launched in 2012 (one tablet per student). Mixed methods, interviews, classroom observations and online surveys have been applied. The findings illustrate new forms of teaching practices. Studying technology integration from the angle of a socio-technical-pedagogical practice, it reveals the interrelationship of teaching processes and quality of learning. This study shows five forms of Digital Didactical Design in practice, which affect tablet-mediated learning expeditions – most of the designs boost learning, others restrict learning.

**Keywords:** Media tablets, didactical design, empirical study, digital didactics.

## 1 Introduction

Over last years, the increase of mobile devies in classrooms exploded. Almost every student has a smartphone; the mobile devices do not stop at the door of the classroom [12]. It raises a pressure on how to handle these technologies in teaching [40].

In earlier times, Information and Communication Technology (ICT) was segregated from the normal teaching classroom, for instance, in computer labs. This has been changed with the invention of smaller flexible devices, like media tablets. There is a shift from separating ICT and education to *co-located settings* [10]; mobile technology becomes *part* of classrooms; both merged into new communication spaces. In addition, there is the change that teachers perceive the mobile device more positive than years ago [16]. For teachers, media tablets are not only a version of a lightweight laptop [23]. For them, it is a new kind of technology that combines several features of both laptops and handheld devices and became a rather new multimodal device in teaching practices [17] [18]. Research on mobile technologies in K-12 education reports opportunities for improving student engagement and achievement of learning aims [7]. It can be assumed that the new situations affect teaching in different aspects.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 165–178, 2014.
© Springer International Publishing Switzerland 2014

However, "technology will probably not change what it takes to learn (…) but it may change how the process of learning is facilitated" [28] – the adoption matters. Kaganer et al. [24] showed that media tablets improve learning when "truly integrated" into learning settings. The question remains what truly integrated means regarding the teaching practice.

In this study, we look on how teachers meet the new challenges, how they integrate the media tablets. More specifically, from a European Digital Didactics approach, the study asks how the new situation affects *digital didactical designs* in co-located settings where physical teaching spaces and web-enabled tablets *merge into multi-overlaid co-located communication spaces*. In Scandinavian schools, which launched a one-to-one program, one media tablet per child, the study explores the richness and diverse pictures of teachers' didactical designs applied in a media tablet rich environment. Based on empirical data, our study reflects on designs-in-practice and helps educators to make informed choices and decisions about engaged classrooms.

## 2 Theoretical Framework

Studies on media tablets in education illustrate that mobile devices are useful for creating content in an interactive way [34] [15]. Such devices improve teaching practices [7] that shift towards learning-centred classrooms [31] [36]. International studies reveals that such devices create a new quality for open access to information [21], useful for user-generated contexts [37], and are able to change the ownership and power relations [40]. Studies also show that the use of mobile ICT can enable student creativity and student collaboration [8] [9]. However, highly relevant is whether the tablets are integrated into the pedagogical design or not. A pure focus on the media tablets cannot explain the emergence of new digital didactical designs "since tools are always specific to tasks" ([41], p. 155). Research on the integration of technology, pedagogical and content/subject knowledge by teachers, known as TPaCK models [27] [29] points toward a lack of those studies. They focus on micro levels of learning from a learner's perspective; they neglect that teaching is also a *design project* developed and carried by teachers. On the other hand, "the discourse surrounding TPCK is in our view largely blackboxing how teachers appropriate, coordinate and collaborate through educational technology at the level of *practice*" [17]. Still under-explored and under-researched is the teacher's practice, how s/he enables learner-centred classrooms in co-located communication spaces.

Innovations such as new technologies lead to a new situation in schools on different levels. The use of media tablets affects many layers of education, how humans act in the classroom to the content, and affects different layers of European *Didaktik*: 1) the relation between teacher, students and content – interaction model (inner layer), 2) the didactical design (teaching aims, learning activities, social relations and process-based assessment) (middle layer) and 3) the didactical conditions including curriculum development (including examination), strategic institutional management and academic staff development (outer layer), figure in [17]. We focus in this paper on the middle layer, the Digital Didactical Design.

***Digital Didactical Design (DDD) – act of modeling and forming the teaching practice for learner-centred classrooms.*** We define teaching practices as sociotechnical-pedagogical processes where the enablement of learning is the central purpose. It is not possible to deliver learning like people deliver products. However, it is possible to create opportunities for enabling learning. Teachers can restrict or enable learning by applying different designs helping to increase the likelihood that learning takes place [42]. This understanding has its foundation in the learning-centered paradigm, where students construct meaning rather than teachers deliver knowledge [11].

When European Didaktik (definition in next paragraph) claims that a didactical design is the foundation for planning and doing the teaching practice towards a shift from teaching-centered to learning-centred approaches [32] [2] [42], then research should be able to study such learner-centered applied didactical designs in practice.

The term "Didactics" (didactical) comes from the Scandinavian and German concept of Didaktik [26] and focuses on the relation between content-student-teacher, and stress the differences of teaching activities and learning activities ([30], [14]). Didaktik does not only include methods, 'how' to teach, but also embraces the question of 'what' to learn (curriculum and content), 'why', and 'when/where', in what kinds of situations and locations, and how it can be reached, e.g., resources, institutional and academic staff development. One central component in Didaktik is the cultivation of social relationships. Without this, a didactical design would be mainly teacher-led instead of learner-centred.

We call it "digital" since in an Internet-driven world, teaching practices are always technology-based; but they do range from low to high extent of use supporting different forms of learning where the quantity and the quality of the technology-integration vary. We focus on media tablets, instead tabletPCs, since the latter also represents a laptop where the screen is used as tablet, whereas a media tablet is a rather a small flexible lightweight device with easy Internet access via WiFi and 3G/4G.

With the concept of "design", the focus lies on specific educational components. A design is the act of giving a *form*; it shapes a focus and key points for doing teaching; it is process and product at once. A design makes specific teacher's actions and activities visible. It put certain elements of a classroom in the center but does not take the whole reality into consideration. A design is the teacher's act of modeling the teaching practices with the purpose to enable engaged student learning.

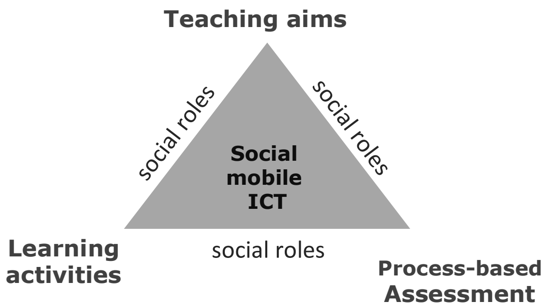



**Fig. 1.** Teacher's Design Elements and Relations – Framing Digital Didactical Design (DDD)

In *an ideal/dream world*, a teacher aligns following components to create a *form* for a learner-centred approach:

- teaching aims and intended learning outcomes are clear and visible
- learning activities are focused mainly on deeper learning (definition below)
- assessment is process-based (informed by guided reflections and "networked scaffolding" ([39], [4]), e.g., 3x milestones to adjust the DDD dynamically
- support of social relations and multiply social roles [19] [22]; teachers are process-mentor/learning-companion and students are pro-sumers/co-designers
- web-enabled mobile technology is multimodal integrated and gives access to overlaid, co-located communication spaces

How do the teachers do this in their practice; what *forms* do they give to their teaching practice; what designs do they apply in the classroom? The five elements together build the frame for modeling the Digital Didactical Design in practice (Figure 1); it links teaching processes to learning. The creation of such an environment affects whether learning can be achieved or not; when these five elements are constructively aligned, they build a purposeful *form*, then the likelihood is high that learning takes places and students are able to learn and meet the intended learning outcomes. A "constructive alignment" [5] is like a puzzle where the pieces complement each other to show a bigger picture.

This view on didactics, design and alignment put studies on technology-enhanced teaching and learning into a new light. Learning is not only a cognitive effort and teaching is not a delivery activity to reach the cognitive dimension. Instead, teaching is rather an activity-driven design, and learning is an on-going activity of knowledge production instead of consumption.

***Quality of Learning – Continuum from Surface to Deeper Learning.*** Following the constructivism approach, learning is knowledge co-construction defined as co-creation of new knowledge that is "an active process of constructing rather than acquiring knowledge" ([11], p. 171). Active learning is related to the role of the learners, where they are not only consumers of information but also active agents and producers in the co-construction of new knowledge: pro-sumers.

Based on Kember's study [25], teaching has been seen for many years as delivery activity *as if* one could deliver learning where textbook readings are in the center of learning. Surface learning is described as remembering facts [1] and teaching contributes to surface learning in supporting the repetition of 'what is in the textbook' (teacher-centred classrooms). Deep learning approaches include activities such as evaluating, creating multi-perspectives [20], collaborative reflections [38]. Teaching that supports both surface and deeper learning helps the learners expand their thinking beyond consumptive behavior and beyond traditional reproduction of existing knowledge (intellectual, "conceptual change" [25]): the learner-centred classroom.

To contribute towards a matrix of quality of learning, a differentiation between individual and group learning is useful, too. The different *designs-in-practice* reach from surface, individual to deep, collaborative learning. Table 1 shows the continuum presented in table. Kember's study does not show how to support the move from surface learinng SI to DG deeper learning (illustrated by the arrow).

**Table 1.** Quality of Learning (Does the Teaching Practice mainly focus on SI, SG, DI or DG?)

| | **Individual learning (I)** | **Learning in groups (G)** |
|---|---|---|
| Deeper com-<br>lex learning<br>(D) | *Examples (DI)*<br>Multimodality,<br>Critical thinking,<br>Analyzing something | *Examples (DG)*<br>Peer-reflective learning<br>Evaluating and Creating in teams<br>*(producer role)*<br>*learner-centred classrooms* |
| Surface learn-<br>ing (S) | *Examples (SI)*<br>Text-book readings<br>Remembering, Understanding<br>*(consumer role)*<br>*teacher-centred approaches* | *Examples (SG)*<br>Group learning:<br>students split the tasks<br>Applying |

Our study explores the digital didactical designs applied by teachers in tablet-classrooms with regard to the quality of learning. The main research question is: What *forms* of digital didactical designs do teachers apply in their media tablet classrooms to support what kind of learning quality; to what extent and how? (RQ).

## 3 Methods

The community of Odder in Denmark with about 20,000 inhabitants and 7 schools implemented media tablets for all of their around 170 teachers and 2,000 students (K-9). The students got the media tablets in a 1:1 program (each child one media tablet) in January 2012. The municipality decided to buy media tablets because their laptops were out-dated. Instead of using new laptops, the community decided to use iPads. The school leaders and the local department of the teachers union were consulted and all parties agreed to the media tablets.

A qualitative approach with mixed methods [6] has been applied, particularly, classroom observations, teacher interviews, school visits (usually 1 school per day) and meetings with head teachers, as part of a larger study about media tablets and European didactics in schools and universities. In total, the research team was in the schools for 4 weeks (20 days) distributed over 2 years. Based on a voluntary sampling [6], 24 classroom observations (45-90 mins. each) plus interviews with the teachers (ca. 60 mins. each) were conducted in six schools in April 2012, August 2012 and August 2013; 7 male teachers and 17 female teachers. The teaching subjects ranged from Native Language, Math, English, Art, Music to Science such as Physics and Biology. The classes ranged from preschool class to 9$^{th}$ grade with different class sizes of 10 up to 27 students (a mix of male and female students). The *classroom observations* were usually conducted by two to three researchers with a training on how to do it and reflections after. They took notes, photos and video recordings with teacher permission. The classroom observations were theoretical guided and based on the digital didactical design elements, in particular, teaching aims, learning activities, forms of assessment, social relations, and degree of media tablet integration into the learning activities. Based on the experiences of a first round of observations, the observation sheet has been developed further. Through the data, it became clear that the concept of "constructive alignment" of the involved elements is a key issue and the research team started to problematize the role of technology. The data illustrated that there is a differentiation between the extent of the tablet use – from low, medium to high extent. The observation sheet included (1) a description of the classroom from the Digital Didactical Design approach, to what extent the elements are in a constructive alignment, (2) how the media tablets are applied in the classroom, (3) communication patterns, social relations, roles (4) collaboration, forms of cooperative learning, (5) feedback and assessment, (e.g., process-based, when, how) (6) what is good/bad from the observer's point of view and why, (7) student learning, creative aspects, (8) special skills of the teacher, (9) anything else. The *interviews* were conducted by a total of three researchers and recorded. The interview guide was divided into five parts and contained 12 questions guided by the didactical design approach.

Data from the observations and interviews were first analyzed according to each classroom and then open coded [3]. For the data analysis, a scheme has been derived from the Digital Didactical Design model in order to make the different designs-in-practice for each classroom visible. The data were coming from the observations and the interviewed teachers. Each classroom was evaluated on this scheme (Table 2).

**Table 2.** Scheme for data analysis (per classroom)

| **Orientations towards Learner-centred classrooms** | Description - data based | Component addressed? Y/N | **To what extent does the component support the learner-centred classroom?** 5=strong; 1= weak alignment |
|---|---|---|---|
| **Teaching aims** are visible/ clear? Are intended learning outcomes visible/clear? **Co-aims** by students included? | | | *Details of the design* |
| **Learning activities** are clear and appropriate, corresponding to teaching aims? | | | Surface to deep learning? Engaged classrooms? |
| **Feedback**: assessment is process-based? Guided reflections teacher/peers/self-assessm.? | | | Process-based? |
| **Social relations/roles:** teachers/students act in multiply roles? Which ones? Does the teacher explicitly support the cultivation of social relations; if yes, how? | | | Teacher=expert, process mentor, learning-companion, ...? Student=consumers, producers, collaborators, …? |
| **Is the media tablet multimodal integrated?** into the whole learning scenario | | | What is the purpose of using the media tablet; what activities are supported? |
| **Overall analysis per classroom** | Brief summary | How many addressed components in total? | -Summary (how many aligned components in total towards a learner-centred approach?) -Extent of tablet use as high (3) medium (2) low (1) |

***Low, Medium, High Extent of Tablet Use.*** The research team analyzed the teachers' applied designs based on what they did in the classroom including what they said in the interviews. In addition, the extent of the media tablet integration was analyzed: a low, a medium and a high extent. A low extent is defined as a non-value of using media tablets in the classroom situation or it was not evident; for example, the media tablet is a substitute for pen and paper or a textbook substitute. A medium extent is assigned when the media tablets are a substitute for other existing digital devices that also could have been used, for example, a computer, laptop or a digital camera. A high extent is defined when the use of the media tablet shows special characteristics or features what no other device can make right now, for example, special apps, an one-in-all device, a multi-modal device.

The analyzed data were checked by content- and peer-review validation; at least 3 researchers checked the analysis of the data. Such a *communicative* validation was done by using inter-subjective methods that proves the quality of the findings [3].

## 4 Findings

The 24 observed classrooms have been analyzed to make differences and similarities visible. For every classroom, the data can be analyzed towards its applied Digital Didactical Design (Fig. 2); where the inner frame represents a rather teacher-centred (1) and the outer frame represents a learner-centered approach (5); on a 5-point scale from 1 to 5. Figure 2 gives an example of three analyzed classrooms (marked in different shades of grey). To each classroom, a summarized value has been derived, which is shown in Table 3 under "how many elements support the learner-centred classroom?". Table 3 illustrates the total amount in combination with a low, medium or high extent of the tablet use. The data indicates different *forms* of Digital Didactical Design; e.g., the classroom ID 6 has five aligned elements (=5) plus a high extent of the tablet use (=3) makes 8 in total. The research team defined that 1-2 is one *form*, 3-4 is another one, 5-6 is the third form and 7-8 is a fourth form. A special case is ID 3. Five *forms* have been explored named as MD, DD, BT, PD, RE.

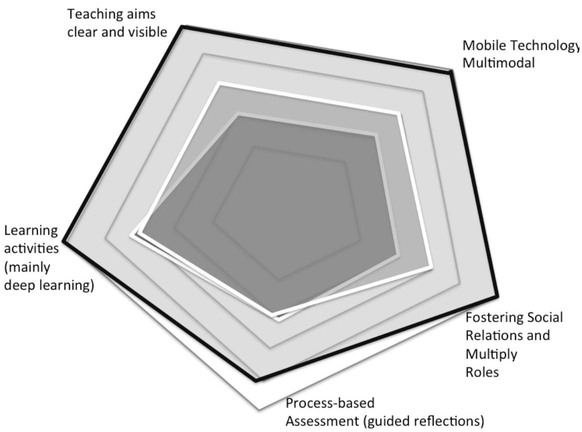



**Fig. 2.** Example of three analyses– one *form* per classroom (inner circle 1=teacher-centered; outer circle 5=learner-centred class)

**Table 3.** Classes representing different *forms* of Digital Didactical Design

| **ID** | **Subject (grade)** | **To what extent do the five elements support the learner-centred classroom and how many are aligned?)** | **Media tablet extent of use** high=3; medium=2, low=1 | **Forms** |
|---|---|---|---|---|
| 5 | Math (2nd) | 5 | +High (3) | =MD |
| 6 | Preschool class | 5 | +High | =MD |
| 10 | Arts (8th) | 5 | +High | =MD |
| 11 | Physics (9th grade) | 5 | +High | =MD |

**Table 3.** (*continued*)

| | | | | |
|---|---|---|---|---|
| 17 | Language (7th) | 5 | +High | =MD |
| 18 | Chemistry (9th) | 5 | +High | =MD |
| 20 | History (5th) | 5 | +High | =MD |
| 21 | Language (4th) | 5 | +High | =MD |
| 23 | Language (5th) | 5 | +High | =MD |
| 4 | Math (1st) | 4 | +High | =MD |
| 24 | Language (9th) | 4 | +High | =MD |
| 8 | Writing skills (7th) | 4 | +Medium (2) | =DD |
| 7 | Music (6th) | 3 | +High | =DD |
| 15 | Geography (3rd) | 3 | +High | =DD |
| 22 | Geography (4th) | 3 | +Medium | =DD |
| 16 | Religion (3rd) | 3 | +High | =DD |
| *3* | *History (2nd)* | *1* | *+High (3)* | *=BT* |
| 2 | Language (2nd) | 2 | +Medium | =PD |
| 12 | Geography (5th) | 2 | +Medium | =PD |
| 1 | English (6th) | 2 | +Medium | =PD |
| 13 | Preschool class (0th= | 3 | +Low (1) | =PD |
| 19 | History (3rd) | 0 | +Medium | =RE |
| 9 | Science, Biology (1st) | 0 | +Low | =RE |
| 14 | Language (3rd) | 1 | +Low | =RE |

MD = Media-tablet-Didactics (7-8)
DD = Digital Didactics (5-6)
BT = Benefit of Tablet integration (special case, 4)
PD = Potential for a digital didactical design (3-4)
RE = RE-alignment required (1-2)

The boxplot analysis in Figure 3 shows a correlation between aligned digital didactical designs and a high extent of the tablet use and vice versa.

It is important to stress that every classroom is rich of information itself and provides a *complexity* of information that cannot be mirrored in numbers such as in Table 3. Please read for further information our detailed qualitative study published in [18]. The purpose of Table 3 is to illustrate that there exist different designs-in-practices which can be clustered in different forms. We've chosen these clusters since it is useful to understand the different designs; the five *forms*

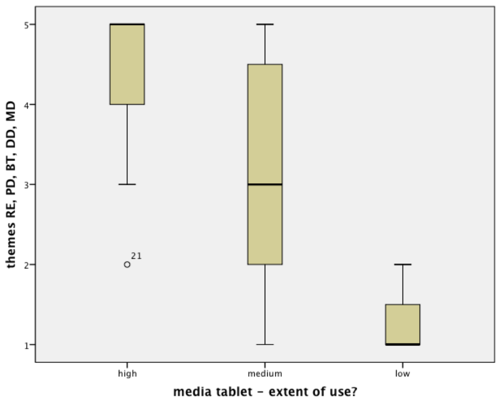



**Fig. 3.** Boxplot analysis - Correlation of digital didactical design forms and extent of tablet use; y-axis: 5=MD, 4=DD; 3=BT, 2=PD, 1=RE

do not go to much into detail (otherwise we would have 24 forms) but also they are not too general (e.g., only 2 clusters); this would be difficult to explore differences/similarities across the cases.

## 5 Discussions and Implications

One possible way of discussing the similarities and differences of classrooms is the exploration of their different applied designs.

*Form of Media tablet Didactics (MD).* 11 of 24 classrooms (ID=4, 5, 6, 10, 11, 17, 18, 20, 21, 23, 24) show the full potential of a digital didactical design in which the elements are constructively aligned to increase the possibilities for learning. The five elements of a DDD are aligned in such a combination that they foster the intended learning outcomes (defined by the teacher) and foster a learner-centred approach.

*Form of Digital Didactics (DD).* There are five classes (ID=7, 8, 15, 16, 22) that used the media tablets as a laptop substitute to reach the intended learning outcomes. The only difference to MD is that the teachers didn't use the unique potential of a media tablet as a multi-modal device. The cases of the DD show that the tablets are also useful when not using its full potential as a multimodal device but more as a laptop function. The teachers said in the interviews, however, when using the tablet like a laptop or for writing assignments, there are some obstacles, for example, there is no keyboard for writing and that makes a tablet slower than a laptop and then, an external keyboard for the tablet is required.

*Form of weak alignment but Benefit of Tablet-integration (BT).* The classroom ID 3 is an interesting case. The class was a traditional class, teacher-led, where process-based feedback and the design of social relations were not aligned at all. The rest of the digital didactical design elements were in a constructive alignment, however, with the goal to foster the traditional teaching such as "Instruction-Response-Feedback" [33] and less collaboration. The observers reported that the whole classroom was rather a weakness of not creating a supporting learning culture, which was obvious during the observation. But the media tablet integration made then the difference. The students got the task to create a movie or a book (students' choice) about the historical person called Kristian IV in order to show the teacher what they have learnt. The collaborative production of a movie by using the iMovie app was an added value to foster learning by producing. Through the phase of producing the students also reflected on what they created and discussed changes.

*Form of Potential for Digital Didactics (PD).* Four cases (ID=1, 2, 12, 13) have in common that the five components of the Digital Didactical Design are rather weak aligned; the added values of the media tablets and why to use them were not clear. The cases also show a weak alignment to reach the intended learning outcomes. The classes did not limit learning but rather did support a non-constructive alignment; a sign for teacher-centred classrooms. The classes integrated the media tablets in a medium to low extent. The potential for a stronger constructive alignment was obvious.

*Form of RE-alignment of a digital didactical design; better without tablets? (RE).* The data reveals three cases (ID=9, 14, 19) in which the integration of media tablets reduced the students learning experiences and restricted learning. The use of the media tablets and the didactical designs in those classes where not connected in such a way it would be beneficial for students learning. Instead it seems that the media tablet was applied in a way that limits the learning activities instead of supporting the intended learning outcomes.

Table 4 illustrates the digital didactical designs and its range from individual to group learning, 13 of 24 classes supported group assignments (9 collaborative deeper learning, 4 group assignments on the surface learning level); 11 classes focused on individual learning where 9 classes did support deeper learning and 2 surface levels. With regard to the MD+DD, 8 out of 16 classes might be labeled as "DG" (deep, group learning) and 0% are in the lower left (SI, surface, individual).

**Table 4.** From surface to deeper learning and from individual to group learning

| | **Individual Learning (I)** | **Learning in Groups (G)** |
|---|---|---|
| **Deeper complex learning (D)** | • Transforming a math story (ID5) MD<br>• Creating a multimodal book review (ID6) MD<br>• Creating chronological order (ID21) MD<br>• Creating a Multimodal Story (ID 16) DD<br>• Finding animals across the globe (ID 22) DD<br>• Individual produc. of audio product (ID 1) PD<br>• Creating a digital story/proverbs (ID 2) PD<br>• Writing a non-fiction story (ID14) RE<br>• Individual timeline (ID 19) RE | • Collaborative writing (ID10) MD<br>• Collaboratively designing of physical experiments (ID11) MD<br>• Creating chemical experiments (ID18) MD<br>• Collaborative creation of a multimodal product (ID20) MD<br>• Collaborative production on "Explains everything" (ID23) MD<br>• Creating a multimodal product from a graphical novel (ID24) MD<br>• Collaboratively producing of music (ID 7) DD<br>• Peer-reflective learning (ID 8) DD<br>• Collaborative production of a video/movie (ID3) BT |
| **Surface learning (S)** | • Role-playing teacher-led (ID 13) PD<br>• Mind mapping existing knowledge (ID9) RE | • Group discussions about math (ID4) MD<br>• Group work outside of the classroom (ID17) MD<br>• Finding/discussing distances, GoogleMaps (ID15) DD<br>• Creating a digital presentation (ID12) PD |

Since the selection of the classroom was based on a voluntary sampling, they are not representative all classrooms in Scandinavia. The data, however, reveals a richness of applied digital didactical designs in the teaching practice. From the perspective of innovative classrooms, the data is useful to analyze the potential of what is possible when using technology such as media tablets and what hinders learning.

***Design Guidance – from Course-Based Learning* to *Learning Expeditions?*** The classrooms illustrate different forms of digital didactical designs (DDD) in practice clustered in five *forms*. It is not a surprise that the usage of media tablets in some of the classrooms focused on enhancing deeper learning and others support surface learning. In some cases the applied design even limited the chance that learning takes place (*form* RE). The merging of all components such as new technology, aims and activities into a new digital didactical design lead to different *forms* – this is what our study illustrates.

When we have a detailed look into the classrooms, which foster group learning, then, one key principle could be explored which they have in common: They created learning opportunities which went from a course-based learning approach to student-centered learning expeditions. The characteristics of – and design guidance for – learning expeditions are:

- *New types of learning goals – more than one correct answer exists*. The teachers in Denmark created those digital didactical designs that enabled learning towards different possible solutions, where no correct answer exists, "learning when the answer is not known" [13].
- *Learning in classrooms moves into design projects*. Adopting the media tablet, the teachers in Odder created digital didactical designs that focus on learning as a *process*. The teachers activated the engagement of the students and their motivation by requesting them *producing* something; knowledge production over consumption. One teacher argued, "*I want to set the knowledge of my students free*" (ID11) that is why he created designs for learning focusing on knowledge production in groups. The media tablets helped to make the student's process visible.
- *From textbook reading to learning that turns into exploring sth*. The teachers combined traditional textbook readings with open-ended, unstructured spaces where students have been encouraged to experiment, play and explore topics.
- *Teachers foster students to make their learning visible in different products (and apps)*. While using different apps, the students shared their learning situations to learn from each other via co-located communication spaces. The assignments were created in a way that the students could choose how to make their learning visible; they did not only choose "to write". The teachers also supported students to create other products like digital paintings, digital stories, comics, movies and podcasts. The teachers used apps, which were originally not made for school purposes (e.g., Bookcreator, Puppetpals, Popplet, Stripdesign, Comicbook).

MD and DD clusters show that the classrooms moved from teacher-centred concepts to *Learning Walkthroughs* that presupposes a rather designed learning landscape more closely guided by teaching but with a greater variation and more student options to work and learn than the traditional course. The next step towards learning expeditions needs still to be done. Some classrooms (e.g., ID18, ID24) are on the way towards *Learning Expeditions* which are similar to learning walkthroughs; however, learning expeditions are more open-ended, problem-based learning paths, and contribute to goal/objective-oriented learning (e.g., to master X or to explore and understand the implications of Y) and the methods and instruments are very open. Learning does not take place in straight-ahead processes but in *loops* and detours, back & forth.

With this approach, we provide an alternative evaluation matrix for educational institutions for their teaching and learning practice; they don't need to rely on PISA only. This alternative option combines a two step approach: (a) studying Digital Didactical Designs in practice, including a qualitative in-depth description of the cases (published in [18]) and (b) the perceived value of learning from the students' perspectives as published in [35] available in this volume of ECTEL2014 proceedings.

## 6 Conclusion

The aim was to explore teaching practices from the perspective of digital didactical designs and to make similarities and differences visible. The study explored five forms of digital didactical designs towards a engaged learner-centred classroom practice from bottom-up, a) specific media-tablet-Didactics, b) digital didactics, c) weak alignment but benefit of media tablet-integration, d) potential for digital didactics and e) re-alignment of designs for learner-centred classrooms. Some teachers created new digital didactical approaches they transformed their traditional classrooms into creative learning expeditions for their students. Other data showed potential for improvements, which the research team discussed with the teachers. Our main findings are:

- We developed a generic language for discussing, exploring and observing teaching and learning in technology rich environments where educational technology is used towards a shift from teacher-led to learner-centred classrooms; we call it Digital Didactical Design (DDD).
- This model was then used to analyze the data in schools in a Scandinavian country. Based on the data, we are also able to revise the DDD model.
- With the DDD model, we are now able to present the richness of different designs in the teaching practice in the context of 1:1 technology programs.
- The DDD model is especially suitable in contexts where new paradigms of teaching and learning in educational institutions such as Learning Expeditions are emerging where new technology is *one part* in this shift. The language of the model may be seen as primarily connected to the mesolevel (teacher's practice and student learning in classrooms); it does not focus on the micro-level in detail (human-content-interactions) or macro-level (institutional strategy viewpoint).
- With the reflection by the help of DDD in combination with the further development of the DDD, we now have an approach that focus on a) teachers' applied designs-in-practice b) technology in relation to the teacher's design, merging into new designs, and c) the extent of technology use. It means that we now have at least three ways to how we are able to present the results from a multimodal perspective and that is one benefit of the DDD model.

The teacher's role is one important engine for innovation; it makes the difference when the Digital Didactical Design (DDD) is aligned towards learner-centred classrooms. The innovative teachers in Odder (DK) applied new forms of DDD. Our study shows how such designs contribute to deeper learning to help the learners to reflect and deepen their skills on their way to become critical-constructive pro-sumers.

**Acknowledgements.** We thank the municipality project leader Lise Gammelby, the school leaders, head teachers and all teachers in Odder, which do an amazing job in being creative and reflective in their use of media tablets in teaching and learning. We are very grateful to be part of this project.

## References

1. Anderson, L.W., Krathwohl, D.R.: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman, NY (2001)
2. Barr, R.B., Tagg, J.: From teaching to learning—A new paradigm for undergraduate education. Change: The Magazine of Higher Learning 27(6), 12–26 (1996)
3. Bauer, M., Gaskell, G.: Qualitative Researching with text, image and sound. Sage, London (2000)
4. Bergström, P.: Designing for the Unknown. Didactical Design for Process-based Assessment in Technology-Rich Learning Environments. Umeå University Press (2012)
5. Biggs, J., Tang, C.: Teaching for Quality Learning at University, 3rd edn., New York (2007)
6. Bryman, A.: Social research methods. Oxford University Press, New York (2008)
7. Chou, C.C., Block, L., Jesness, R.: A case study of mobile learning pilot project in K-12 schools. Journal of Educational Technology Development 5(2), 11–26 (2012)
8. Buchem, I., Jahnke, I., Pachler, N.: Guest editorial preface. SI Mobile Learning and Creativity. International Journal of Mobile and Blended Learning 5(3) (2013)
9. Buchem, I., Cochrane, T., Gordon, A., Keegan, H.: M-Learning 2.0: The potential and challenges of collaborative mobile learning in participatory curriculum development. In: Proceedings of the IADIS Mobile Learning Conference 2012, Berlin, Germany (2012)
10. De Chiara, R., Di Matteo, A., Manno, I., Scarano, V.: CoFFEE: Cooperative Face2Face educational environment. In: Collaborative Computing: Networking, Applications and Worksharing Conference, pp. 243–252 (2007), doi:10.1109/COLCOM.2007.4553836
11. Duffy, T.M., Cunningham, D.J.: Constructivism: Implications for the design and delivery of instruction. In: Handbook of Research for Educational Communications and Technology, pp. 170–199. Simon & Schuster Macmillan, New York (1996)
12. Findahl, O.: Svenskarna och Internet (Sweden and the Internet) (2012), `https://www.iis.se/docs/SOI2012.pdf` (retrieved February 12, 2012)
13. Fischer, G.: Social Creativity: Exploiting the Power of Cultures of Participation. In: Proceedings of SKG 2011: 7th International Conference on Semantics, Knowledge and Grids, Beijing, China, pp. 1–8 (October 2011)
14. Hudson, B.: A Didactical Design Perspective on Teacher Presence in an International Online Learning Community. Journal of Research in Teacher Education 15(3-4), 93–112 (2008)
15. Hutchison, A., Beschorner, B., Schmidt-Crawford, D.: Exploring the Use of the iPad for Literacy Learning in. The Reading Teacher 66(1), 15–23 (2012)
16. Ifenthaler, D., Schweinbenz, V.: The acceptance of Tablet-PCs in classroom instruction: The teachers' perspectives. Computers in Human Behavior 29(3), 525–534 (2013)
17. Jahnke, I., Svendsen, N., Johanson, S., Zander, P.-O.: The Dream About the Magic Silver Bullet – the Complexity of Designing for Tablet-Mediated Learning. In: ACM GROUP Conference Proceedings (2014)
18. Jahnke, I., Kumar, S.: Digital Didactical Designs: Teachers' Integration of iPads for Learning-Centered Processes. Journal of Digital Learning in Teacher Education 30(3), 81–88 (2014), doi:10.1080/21532974.2014.891876
19. Jahnke, I.: Dynamics of social roles in a knowledge management community. Computers in Human Behavior 26 (2010), doi:10.1016/j.chb.2009.08.010
20. Jahnke, I.: How to Foster Creativity in Technology Enhanced Learning. In: White, B., King, I., Tsang, P. (eds.) Social Media Tools and Platforms in Learning Environments, pp. 95–116. Springer, NY (2011), doi:10.1007/978-3-642-20392-3_6
21. Jahnke, I., Bergström, P., Lindwall, K., Marell-Olsson, E., Olsson, A., Paulsson, F., Vinnervik, P.: Understanding, Reflecting and Designing Learning Spaces of Tomorrow. In: Arnedillo Sanchez, I., Isaias, P. (eds.) Proceedings of IADIS Mobile Learning 2012, Berlin, pp. 147–156 (2012)

22. Jahnke, I., Ritterskamp, C., Herrmann, T.: Sociotechnical Roles for Sociotechnical Systems: a perspective from social and computer science. In: AAAi Fall Symposium Proceedings, 8. Symposium, Roles, an Interdisciplinary Perspective, Arlington, Virgina (2005)
23. Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., Ludgate, H.: NMC Horizon Report: 2013 Higher Education. New Media Consortium, Austin (2013)
24. Kaganer, E., Giordano, G.A., Brion, S., Tortoriello, M.: Media Tablets for Mobile Learning: Friend or Foe? Communications of the ACM 56(11), 68–75 (2013)
25. Kember, D.: A reconceptualisation of the research into university academics' conceptions of teaching. Learning and Instruction 7(3), 255–275 (1997)
26. Klafki, W.: Studien zur Bildungstheorie und Didaktik. Beltz, Weinheim (1963)
27. Koehler, M.J., Mishra, P., Yahya, K.: Tracing the development of teacher knowledge in a design seminar: Integrating content, pedagogy and technology. Computers & Education 49, 740–762 (2007)
28. Laurillard, D.: Technology enhanced learning as a tool for pedagogical innovation. Journal of Philosophy of Education 42(3-4), 521–533 (2008)
29. Loveless, A.: Preparing to teach with ICT: subject knowledge, Didaktik and improvisation. Curriculum Journal 18(4), 509–522 (2007), doi:10.1080/09585170701687951
30. Lund, A., Hauge, T.E.: Designs for teaching and learning in technology-rich learning environments. Nordic Journal of Digital Literacy (4), 258–272 (2011)
31. McCombs, S., Liu, Y.: Channeling the channel: Can iPad meet the needs of today's M-Learner. In: Proceedings of Society for Information Technology & Teacher Education, Chesapeake, VA, pp. 522–526 (2011)
32. McCombs, B.L., Whisler, J.S.: The learner-centered classroom and school: Strategies for increasing student motivation and achievement. Jossey-Bass, San Francisco (1997)
33. Mehan, H.: Learning lessons. The social organization of classroom instruction. Harvard University Press, Cambridge (1979)
34. Melhuish, K., Falloon, G.: Looking to the future: M-learning with the iPad. Computers in New Zealand Schools 22(3), 1–16 (2010)
35. Norqvist, L., Jahnke, I., Olsson, A.: The Learners' Expressed Values of Learning in a Media Tablet Learning Culture. In: EC-TEL 2014. LNCS, vol. 8719, pp. 460–465. Springer, Heidelberg (2014)
36. Ostashewski, N., Reid, D.: iPod, iPhone, and now iPad: The evolution of multimedia access in a mobile teaching context. In: Proceedings of World Conference on Educational Multimedia, Hypermedia, pp. 2862–2864. AACE, Chesapeake (2010)
37. Pachler, N., Bachmair, B., Cook, J.: Mobile learning: Structures, agency, practices. Springer, New York (2010), doi:10.1007/978-1-4419-0585-7
38. Prilla, M., Herrmann, T., Degeling, M.: Collaborative reflection for learning at the healthcare workplace. In: Goggins, S.P., Jahnke, I., Wulf, V. (eds.) Computer-Supported Collaborative Learning at the Workplace, pp. 139–165. Springer, NY (2013)
39. Tammets, K., Laanpere, M., Ley, T., Pata, K.: Identifying Problem-Based Scaffolding Patterns in an Online Forum for Construction Professionals. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 526–531. Springer, Heidelberg (2013)
40. Traxler, J.: Context in a wider context. Medienpädagogik. Mobile Learning in Widening Contexts: Concepts and Cases, 19 (2011)
41. Wegerif, R.: A dialogic understanding of the relationship between CSCL and teaching thinking skills. Computer Supported Collaborative Learning 1(2006), 143–157 (2006), doi:10.1007/s11412-006-6840-8
42. Wildt, J.: On the Way from Teaching to Learning by Competences as Learning Outcomes. In: Pausits, A., Pellert, A. (eds.) Higher Education Management and Development in Central, Southern and Eastern Europe, pp. 115–123. Waxmann, Münster (2007)

# Feature Analysis for Affect Recognition Supporting Task Sequencing in Adaptive Intelligent Tutoring Systems

Ruth Janning, Carlotta Schatten, and Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim,
Marienburger Platz 22, 31141 Hildesheim, Germany
{janning,schatten,schmidt-thieme}@ismll.uni-hildesheim.de
http://www.ismll.uni-hildesheim.de

**Abstract.** Originally, the task sequencing in adaptive intelligent tutoring systems needs information gained from expert and domain knowledge as well as information about former performances. In a former work a new efficient task sequencer based on a performance prediction system was presented, which only needs former performance information but not the expensive expert and domain knowledge. This task sequencer uses the output of the performance prediction to sequence the tasks according to the theory of Vygotsky's Zone of Proximal Development. In this paper we aim to support this sequencer by a further automatically to gain information source, namely speech input from the students interacting with the tutoring system. The proposed approach extracts features from students speech data and applies to that features an automatic affect recognition method. The output of the affect recognition method indicates, if the last task was too easy, too hard or appropriate for the student. Hence, as according to Vygotsky's theory the next task should not be too easy or too hard for the student to neither bore nor frustrate him, obviously the output of our proposed affect recognition is suitable to be used as an input for supporting a sequencer based on the theory of Vygotsky's Zone of Proximal Development. Hence, in this paper we (1) propose a new approach for supporting task sequencing by affect recognition, (2) present an analysis of appropriate features for affect recognition extracted from students speech input and (3) show the suitability of the proposed features for affect recognition for supporting task sequencing in adaptive intelligent tutoring systems.

**Keywords:** affect recognition, feature analysis, speech, task sequencing, adaptive intelligent tutoring systems.

## 1 Introduction

Nowadays, intelligent tutoring systems are an important tool for supporting education of students for instance in learning mathematics. The main advantages of intelligent tutoring systems are the possibility for a student to practice any time, as well as the possibility of adaptivity and individualisation for a single

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 179–192, 2014.
© Springer International Publishing Switzerland 2014

student. Usually, an adaptive intelligent tutoring system possesses an internal model of the student and a task sequencer which decides which tasks in which order are shown to the student. Originally, the task sequencing in adaptive intelligent tutoring systems is done using information gained from expert and domain knowledge and logged information about the performance of students in former exercises. In [12] a new efficient sequencer based on a performance prediction system was presented, which only uses former performance information from the students to sequence the tasks and does not need the expensive expert and domain knowledge. This approach applies for performance prediction the machine learning method matrix factorization to former performance information. Subsequently, it uses the output of the performance prediction process to sequence the tasks according to the theory of Vygotsky's Zone of Proximal Development [14]. That is the sequencer chooses the next task in order to neither bore nor frustrate the student or in other words, the next task should not be too easy or too hard for the student. In this paper we propose to support and improve the sequencing approach in [12] by a further automatically to get and process information, namely speech input from the students interacting with the system while solving tasks. The proposed approach extracts from students speech data features which are adapted for supporting task sequencing and applies to that features an automatic affect recognition method. The output of the affect recognition method indicates, if the last task was too easy, too hard or appropriate for the student. Obviously, this output can be used as an input for supporting a sequencer based on the theory of Vygotsky's Zone of Proximal Development. To the best of our knowledge, supporting task sequencing by this kind of features gained from speech was not done before, and in this work we will show a promising analysis of such features. The main contributions of this paper are: (1) a proposal and explanation of a new approach for supporting task sequencing by affect recognition, (2) presentation and analysis of own appropriate features extracted from students speech input of a real data set, (3) verification of the possibility to use the proposed features for affect recognition for supporting task sequencing in adaptive intelligent tutoring systems. In the following sections first we will present some preliminary considerations about appropriate speech features, affect classes, classification instances and possible methods for affect recognition supporting task sequencing along with state-of-the-art (sec. 2), followed by a short description of the investigated real data set (sec. 3). Subsequently, we will describe in section 4 our proposed features and their analysis. Before we conclude we will explain in section 5 how to realise the desired support of task sequencing by affect recognition.

## 2 Preparation and Related Work

As mentioned above, we propose to apply an automatic affect recognition approach to students speech input, whose output shall serve as input for supporting a task sequencer based on the theory of Vygotsky's Zone of Proximal Development. However, before an automatic affect recognition approach can be applied, one has to clarify three things: (1) What kind of features shall be used, (2) what kind of classes shall be used and (3) which instances shall be mapped to features and labelled with the class labels. After deciding which features, classes and instances shall be considered, one can apply affect recognition methods to these input data. In the following subsections we will present possible features, classes, instances and methods for affect recognition supporting task sequencing along with the state-of-the-art.

### 2.1 Features

The first step before applying automatic affect recognition is to identify useful features for this process. For the purpose to recognise affect in speech one can use two different kinds of features (see [13]): acoustic and linguistic features. Further, one can distinct linguistics (like n-grams and bag-of-words) and disfluencies (like pauses). If linguistics features are used, a transcription or speech recognition process has to be applied to the speech input before affect recognition can be conducted. Subsequently, approaches from the field of sentiment classification or opinion mining (see e.g. [10]) can be applied to the output of this process. However, the methods of this field have to be adjusted to be applicable to speech instead of written statements, for instance they have to be robust against speech recognition errors. Another possibility for features is to use disfluencies features like it was done in [17], [8] and [5] for expert identification. The advantage of using such features is that instead of a full transcription or speech recognition approach only for instance a pause identification has to be applied before computing the features. That means that – in the case of speech recognition – one does not inherit the error of the full speech recognition approach. Furthermore, these features are independent from the need that students use words related to affects. For using this kind of features one has to investigate which particular features are suitable for the special task of affect classification in adaptive intelligent tutoring systems. Appropriate features are proposed and an analysis of them is presented in this work in section 4.

### 2.2 Classes

The second step before applying automatic affect recognition is to define the classes corresponding to affective states, which shall be recognised by the used approach. According to [7], [6] and [16] it is possible to recognise in intelligent tutoring systems students affects like for instance confusion, frustration, boredom and flow. As mentioned above, we want to use the students behaviour information gained from speech for supporting sequencing of tasks. In [4] and [12] the theory of Vygotsky's Zone of Proximal Development ([14]) was used for sequencing to keep the student in flow. That means that the goal is to neither bore the student with too easy tasks nor to frustrate him with too hard tasks, but to keep him in the Zone of Proximal Development. Accordingly, we want to use the output of the automatic affect recognition to get an answer to the question "Was this task too easy, too hard or appropriate for the student?". With other words, we want to find out if the student felt under-challenged, over-challenged or as to be in a flow. However, the mapping between confusion, frustration, boredom and under-challenged, over-challenged is not unambiguous as one can infer e.g. from the studies mentioned in [16]. Hence, we will use instead of the seemingly obvious affect classes boredom, frustration and flow three other classes for supporting sequencing by automatic affect recognition: under-challenged, over-challenged and flow. These classes could be summarised as *perceived task-difficulty* classes as we want to recognise the personalised task-difficulty from the view of the student, i.e. if the student felt under- or over-challenged, or as to be in a flow.

### 2.3 Instances

The third step before applying automatic affect recognition is deciding which instances shall be mapped to features and labelled with the class labels. For an affect recognition, which aims at giving motivation or hints to the students, like e.g. in [16], the instances according to speech input can be for instance utterances. For supporting task sequencing instead we need at the end of a task the information if the task overall was too easy, too hard or appropriate for the student, because by means of that information the next task shown to the student shall be chosen according to Vygotsky's Zone of Proximal Development. Hence, an instance for affect recognition for supporting task sequencing has to be the whole speech input of a student for one task.

### 2.4 Methods

The possible methods for automatic affect recognition depend on the kind of features used as input. As mentioned above, we distinct two kinds of features: linguistics features and disfluencies. Linguistics features are gained by a preceding speech recognition process and can be processed by methods coming from the areas sentiment classification and opinion mining ([10]). Especially methods from the field of opinion mining on microposts seem to be appropriate if linguistics features are considered. State-of-the-art methods in opinion mining on microposts use methods based on optimisation approaches (see [2]) and Naive Bayes (see [11]). The process of gaining disfluencies like pauses is different to the full speech recognition process. For extracting for instance pauses usually an energy threshold on the decibel scale is used as in [5] or an SVM is applied for pause classification on acoustic features as in [9]. Appropriate state-of-the-art methods for automatic affect recognition on disfluencies features are – as proposed in [13] and [7] – classification methods like artificial neural networks, SVMs, decision trees or ensembles of those.

## 3 Real Data Set

After identifying features, classes, instances and methods like above one can collect data for a concrete feature analysis and later on for a training of the chosen affect classification method. For enabling a feature analysis we conducted a study in which the speech of 10 German students (10 to 12 years old) and their actions as well as students perceived task-difficulty labels (see tab. 3) were reported. The labelling of these data was done on the one hand concurrently by the tutor and on the other hand retrospectively by a second reviewer. Furthermore, a labelling per exercise (overall 36 consisting of several subtasks) and an overall labelling per student as an aggregation of the labels per exercise was done. During the experiment a paper sheet with fraction tasks was shown to the students and they were asked to paint (with the software Paint) and explain their observations and answers, or to think aloud respectively. We made a screen recording to record the painting of the students and an acoustic recording to record the speech of the students. The screen recoding was only used for the retrospective annotation. The speech recordings shall be used as input for the affect recognition. However, as mentioned above, before applying an affect recognition method we have to identify useful features. Because of the mentioned advantages of disfluencies features we decided to use features gained from speech pauses. Furthermore, speech pauses seem to carry useful information in conjunction with the behaviour of students which are trying to solve tasks. We observed in our study for instance a tendency of over-challenged students to stay longer in silent thinking phases than students in flow. Hence, in the following section we will propose and analyse possible features extracted from speech pauses information, which are suitable for supporting task sequencing.

**Fig. 1.** Graphic of the decibel scale of an example sound file of a student. The two straight horizontal lines indicate the threshold.

## 4 Feature Analysis

The features we will propose and analyse in the following subsections are gained from speech pauses. Hence, first one has to identify pauses within the speech input data. The most easy way to do this is to define a threshold on the decibel scale as done e.g. in [5]. For our preliminary study of the data we also used such a threshold, which we adjusted by hand. More explicitly, we extracted the amplitudes of the sound files, computed the decibel values and generated a graphic of it like the one in fig. 1. Subsequently, we investigated which decibel values belong to speech and which ones to pauses to create from this information an appropriate threshold. In larger data and in the application phase later on, one has to learn automatically the distinction between speech and pauses by either learn a threshold or train an SVM which classifies speech and pauses.

### 4.1 Preprocessing

Before introducing the features which we want to investigate, we have to define some measurements (see tab. 1). Our real data set exists of acoustic recordings from $m$ students, each of which saw $n_{t_i}$ tasks and solved $n_{c_i}$ tasks correctly. The overall score of a student $i$ in this case is the number of correctly solved tasks $n_{c_i}$ divided by the number of seen tasks $n_{t_i}$. After applying the above mentioned decibel threshold to the data, we get for each student $i$ the total length of pauses $p_i$ and the total length of speech $s_i$ within his acoustic recoding. Furthermore, we can count connected pause and speech segments to get the number of pause segments $n_{p_i}$ and speech segments $n_{s_i}$ of a student $i$. The $x$th pause segment is then $p_i^{(x)}$ and the $y$th speech segment $s_i^{(y)}$. By means of these measurements and their combination we can create a set of useful features, which will be presented in the next subsection.

**Table 1.** Measurements from the input data of our study, which are needed for extracting and analysing the desired features

| Symbol | Explanation |
|---|---|
| $m$ | Number of students |
| $p_i$ | Total length of pauses of student $i$ |
| $s_i$ | Total length of speech of student $i$ |
| $n_{p_i}$ | Number of pause segments of student $i$ |
| $n_{s_i}$ | Number of speech segments of student $i$ |
| $p_i^{(x)}$ | $x$th pause segment of student $i$ |
| $s_i^{(y)}$ | $y$th speech segment of student $i$ |
| $n_{t_i}$ | Number of tasks shown to student $i$ |
| $n_{c_i}$ | Number of correctly solved tasks by student $i$ |
| $score$ | Overall score for student $i$ ($\frac{n_{c_i}}{n_{t_i}}$) |

### 4.2 Analysis of Single Features

Our proposed features which we will describe and investigate in the following are listed in table 2. The ratio between the total length of pauses and the total length of speech indicates, if one of them is notable larger than the other one, i.e. if the student was more often in a silent thinking phase or in a speech flow. The frequency of speech and pause segment changes indicates for instance, if there are many short speech and pauses segments or just a few large ones. From the percentage of pauses we can see if the total pause length was much larger

**Table 2.** Features created from speech pauses information

| Feature | Formula |
|---|---|
| Ratio between pauses and speech | $\frac{p_i}{s_i}$ |
| Frequency of speech pause changes | $\frac{n_{p_i}+n_{s_i}}{\max_j(n_{p_j}+n_{s_j})}$ |
| Percentage of pauses of input speech data | $\frac{p_i}{(p_i+s_i)}$ |
| Length of maximal pause segment | $\max_x(p_i^{(x)})$ |
| Length of average pause segment | $\frac{\sum_x p_i^{(x)}}{n_{p_i}}$ |
| Length of maximal speech segment | $\max_y(s_i^{(y)})$ |
| Length of average speech segment | $\frac{\sum_y s_i^{(y)}}{n_{s_i}}$ |
| Average number of seconds needed per task | $\frac{(p_i+s_i)}{n_{t_i}}$ |

than the total speech length, i.e. the student did not speak much but was more thinking silently. The length of maximal pause or speech segment indicates if there was e.g. a very long pause segment where the student was thinking silently or a very long speech segment where the student was in a speech flow. The length of average pause or speech segment gives us an idea of how long on average the student was in a silent thinking phase or a speech flow. The average number of seconds needed per task indicates how long a student on average needed for solving a task. We have to mention that the above described measurements (tab. 1) are related to the overall speech input and performance of a student for all tasks. However, for an application of affect recognition for supporting sequencing later on the features have to be taken from speech input and performance of one single task. The measurements can easily be adapted to single tasks. The only huge difference is the computation of the score. In the case of single tasks, the score can depend on the number (#) of requested hints and of incorrect inputs as proposed in [15]:

$$1 - \left( \frac{\#\text{ hints shown for the task}}{\#\text{ all hints of the task}} + (\#\text{ incorr. inputs for the task} \cdot 0.1) \right) . \quad (1)$$

The meaning behind formula (1) is that each wrong input is punished with a factor of 0.1 and every request of a hint is punished with a factor of one divided by the number of all hints available for the task, so that if every hint was seen then the score will be 0. Except for *Average number of seconds needed per task* every proposed feature in table 2 is also working with the measurements related to single tasks. In the following we will analyse our features with the measurements related to the overall speech input and performance of a student for all tasks, as we want to show the general correlation between the features and students overall affects, or perceived task-difficulty respectively. For our analysis we encoded the overall perceived task-difficulty labels by values to be able to do statistical tests. The encoding is shown in table 3.

We investigated as a first step the relevance of each of the eight features alone by mapping the feature values to the labels, doing a linear regression and

**Table 3.** Encoding of perceived task-difficulty labels by values

| Label | Value |
|---|---|
| under-challenged | 4 |
| flow/under-challenged | 3 |
| flow | 2 |
| over-challenged/flow | 1 |
| over-challenged | 0 |

**Table 4.** p-value, $R^2$ and Adjusted $R^2$ for the three best single features (first three rows) and the best combinations of features (with a p-value smaller than 0.05)

| # | **Features** | **p-value** | $R^2$ | **Adjusted** $R^2$ |
|---|---|---|---|---|
| 1 | Maximal pause | 0.0678 | 0.3577 | 0.2774 |
| 1 | Average length of speech | 0.0873 | 0.3217 | 0.2369 |
| 1 | Percentage of pauses | 0.0923 | 0.3136 | 0.2278 |
| 5 | Ratio pause speech, frequency of changes, seconds per task, average length of pause, average length of speech | 0.0284 | 0.9158 | 0.8106 |
| 5 | Frequency of changes, seconds per task, average length of pause, maximal speech, average length of speech | 0.0305 | 0.9127 | 0.8035 |
| 4 | Ratio pause speech, frequency of changes, average length of pause, average length of speech | 0.0154 | 0.8818 | 0.7872 |
| 4 | Frequency of changes, average length of pause, maximal speech, average length of speech | 0.0272 | 0.8501 | 0.7302 |
| 4 | Ratio pause speech, frequency of changes, seconds per task, average length of speech | 0.0288 | 0.8465 | 0.7236 |
| 4 | Frequency of changes, seconds per task, maximal speech, average length of speech | 0.0354 | 0.8324 | 0.6984 |
| 4 | Frequency of changes, seconds per task, average length of pause, average length of speech | 0.0420 | 0.8199 | 0.6759 |
| 3 | Ratio pause speech, frequency of changes, average length of speech | **0.0117** | 0.8207 | 0.7311 |
| 3 | Frequency of changes, average length of pause, average length of speech | 0.0175 | 0.7944 | 0.6916 |
| 3 | Frequency of changes, maximal speech, average length of speech | 0.0242 | 0.7699 | 0.6549 |
| 2 | Frequency of changes, average length of speech | 0.0327 | 0.6238 | 0.5163 |

measuring the p-value, indicating the statistical significance, as well as the $R^2$ and Adjusted $R^2$ value, indicating how well the regression line can approximate the real data points. However, as expected, single features are not very significant. The feature with the best values for p-value, $R^2$ and Adjusted $R^2$ is *Length of maximal pause segment* with a p-value = 0.0678, $R^2 = 0.3577$ and Adjusted

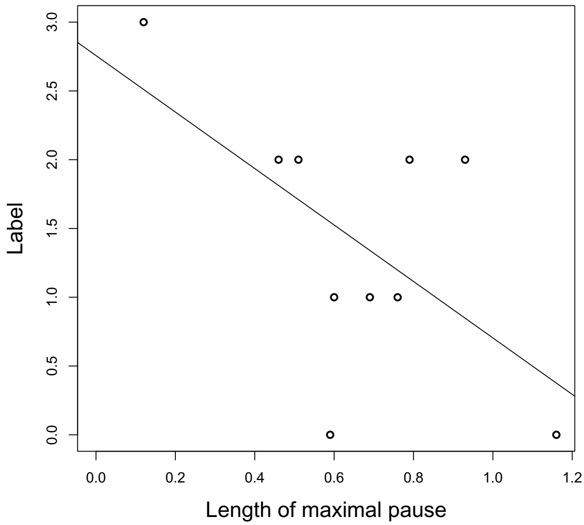



**Fig. 2.** Values of the feature *Length of maximal pause segment* mapped to the labels. Although one can see a tendency in the graphic, unfortunately the p-value, $R^2$ and Adjusted $R^2$ are not very significant.

$R^2 = 0.2774$ (see fig. 2 and tab. 4). In figure 2 one can see that there is a tendency of the values of *Length of maximal pause segment* to increase with decreasing labels. Unfortunately, the values for p-value, $R^2$ and Adjusted $R^2$ of *Length of maximal pause segment* are too weak, as for the p-value a value smaller than 0.05 would be desirable and for $R^2$ and Adjusted $R^2$ a value closer to 1. A better way to use the features is to use a combination of them. Hence, in the next subsection we will investigate feature combinations.

### 4.3 Analysis of Combinations of Features

We investigated different combinations of our features and applied a multivariate linear regression to gain the p-value, $R^2$ and Adjusted $R^2$ of the combinations. The investigated combinations are combinations where all features are not strongly correlated, i.e. whenever we had two correlated features we put just one of them into the feature set for that combination. In further steps we removed from the considered feature sets feature by feature. The best combinations, i.e. such feature combinations with a p-value smaller than 0.05, are listed in table 4, ordered according to the number of features in that set (column # in tab. 4). The feature combination with the best p-value, i.e. which is the most statistically significant one, is a combination of three features (*ratio between pauses and speech*, *frequency of speech pause changes*, *Length of maximal pause segment*) with a p-value of 0.0117, $R^2 = 0.8207$ and Adjusted $R^2 = 0.7311$. These values are strong enough to assume that this combination of features is statistically significant and hence is able to describe students affect, or perceived task-difficulty respectively. That means that it is possible to recognise the perceived task-difficulty of a student by means of these features.

### 4.4 Histograms as Features

In the last section we showed that our features are suitable to describe perceived task difficulty classes. As some of these features cover certain aspects of the histograms of pause segments and of speech segments, an enhancement of the features would be to use the appropriate histograms as features, which deliver more fine granulated information but may increase computation costs. Four example pause and speech histograms of the whole speech input of two students are shown in figure 4. The histograms at the top come from a student with an overall label *over-challenged*, whereas the other histograms at the bottom belong to a student whose behaviour overall was labelled as *flow*. One can see some differences in the graphics, as e.g. the over-challenged student made several very large speech pauses, whereas the student in flow made many very small pauses. On the other hand the student in flow had some very long speech phases, whereas the longest speech phase of the over-challenged student is much smaller than the longest speech phase of the student in flow.

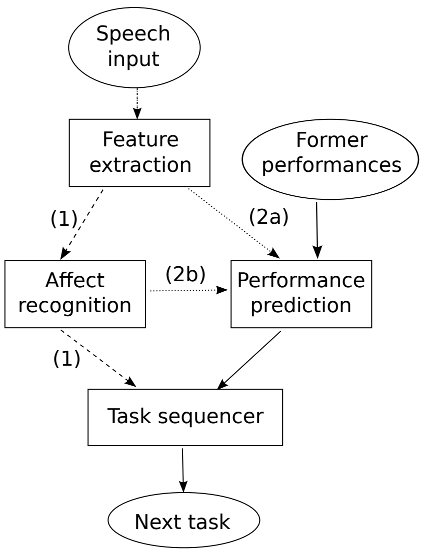



**Fig. 3.** Approach for supporting task sequencing by affect recognition

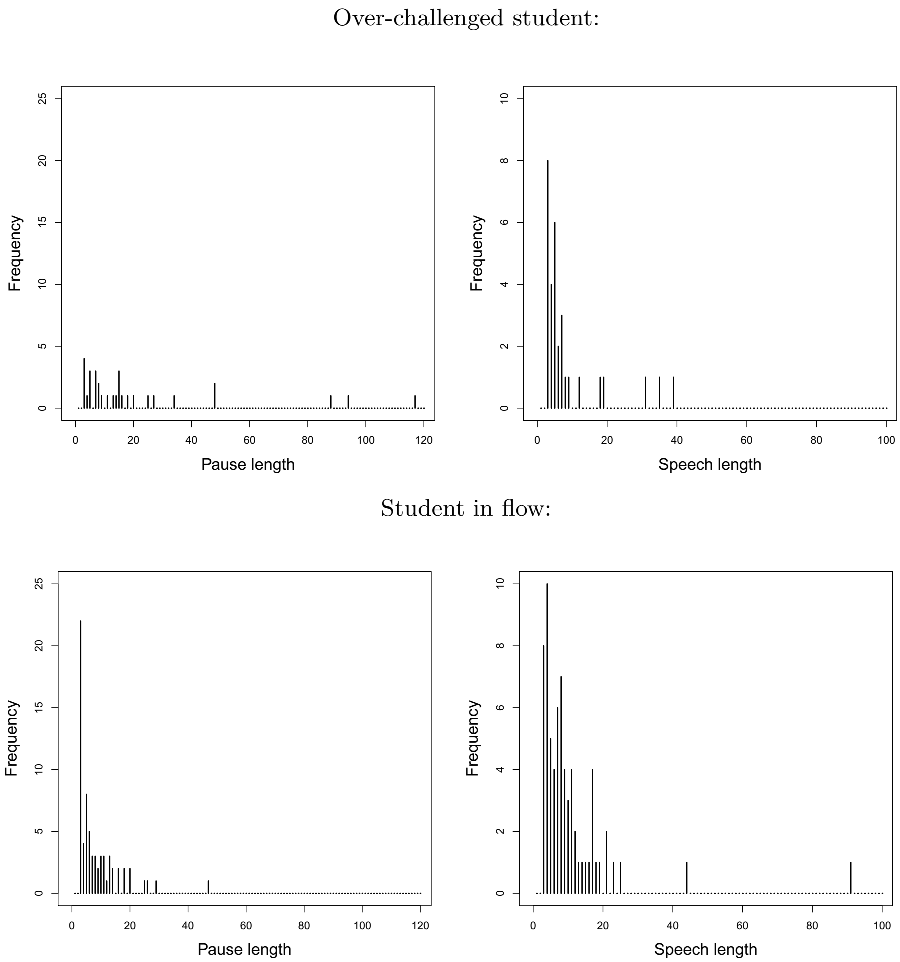



**Fig. 4.** Pause and speech histograms of the whole speech input of two students, one of them was labelled as *over-challenged* (top), the other one was labelled as to be in *flow* (bottom)

## 5 Supporting Task Sequencing in an Adaptive Intelligent Tutoring System

As mentioned in the beginning, our goal is to support task sequencing by affect recognition. More explicitly, we want to support task sequencing based on Vygotsky's Zone of Proximal Development as presented in [12]. The approach in [12] predicts by means of matrix factorization (see e.g. [1]) from former performances the performance of a student for future tasks. Subsequently, the sequencer chooses as next task the task with a predicted performance closest to a certain threshold. This threshold is based on Vygotsky's theory that to keep the student in flow, one should neither bore nor frustrate him. That is this threshold symbolises the performance score which the student would have, if he is in the Zone of Proximal Development. The approach is denoted in figure 3 by the non-dotted arrows. Affect recognition can support this task sequencing approach in several ways (see also fig. 3):

(1) The output of the affect recognition gives us the information, if the last task was too easy, too hard or appropriate for the student. So it indicates if the next task shall be
   (a) more difficult, i.e. the sequencer should choose a task with a smaller predicted performance $\widehat{score}^{(t+1)}$ than the former performance $score^{(t)}$ ($\widehat{score}^{(t+1)} < score^{(t)}$),
   (b) easier, i.e. the sequencer should choose a task with a larger predicted performance $\widehat{score}^{(t+1)}$ than the former performance $score^{(t)}$ ($\widehat{score}^{(t+1)} > score^{(t)}$),
   (c) of similar difficulty, i.e. the sequencer should choose a task with a similar predicted performance $\widehat{score}^{(t+1)}$ as the former performance $score^{(t)}$ ($\widehat{score}^{(t+1)} \approx score^{(t)}$).

   In this case of support the speech features proposed in section 4 are fed into the affect recognition method and the output – the corresponding affect, or perceived task-difficulty class respectively – serves besides the output of the performance prediction as input for the task sequencer (see the dotted arrows with label '(1)' in fig. 3).
(2) The affect recognition can serve as additional input for the performance predictor: (a) the features extracted from speech could serve directly as input for the performance predictor (see the dotted arrow with label '(2a)' in fig. 3), or (b) the affect recognition output could be fed into the performance predictor (see the dotted arrow with label '(2b)' in fig. 3) and let it predict perceived task-difficulty class labels directly, e.g. by a matrix factorization on former perceived task-difficulties.
(3) The affect recognition could help to learn a personalised value for the mentioned threshold based on Vygotsky's Zone of Proximal Development, for instance (a) by considering which performance scores similar students reached on average when they were on flow, or (b) by considering which performance scores the student himself reached on average in former tasks when he was on flow, or (c) by combining both information.

The described support shall be realised within an adaptive intelligent tutoring system, using the mentioned task sequencer based on Vygotsky's theory, which aims at teaching children of the age 8 to 12 at home in front of their computer fractional arithmetic. This adaptive intelligent tutoring system will be the outcome of the EU project iTalk2Learn ([3]).

## 6 Conclusions and Future Work

We proposed a new approach for supporting task sequencing in adaptive intelligent tutoring systems by affect recognition on features gained from students speech input. For this approach we proposed and analysed appropriate speech features and showed that there are statistically significant feature combinations which are able to describe students affect, or perceived task-difficulty respectively. Next steps will be to conduct more studies with students to gain more data for training with the proposed features an automatic affect recognition method, which subsequently can be tested in combination with the sequencer in the mentioned adaptive intelligent tutoring system in further studies. Furthermore, we plan to investigate the combination of our features with results from automatic speech recognition both, on word level as well as on the level of phone-classes such as employed by automatic segmentation.

**Acknowledgements.** This work is co-funded by the EU project iTalk2Learn ([3]) with grant agreement No. 318051.

## References

1. Cichocki, A., Zdunek, R., Phan, A.H., Amari, S.I.: Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation. Wiley (2009)
2. Hu, X., Tang, L., Tang, J., Liu, H.: Exploiting Social Relations for Sentiment Analysis in Microblogging. In: Proceedings of the Sixth ACM WSDM Conference (WSDM 2013) (2013)
3. Talk, Tutor, Explore, Learn: Intelligent Tutoring and Exploration for Robust Learning (iTalk2Learn), `http://www.italk2learn.eu`
4. Krohn-Grimberghe, A., Busche, A., Nanopoulos, A., Schmidt-Thieme, L.: Active learning for technology enhanced learning. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 512–518. Springer, Heidelberg (2011)
5. Luz, S.: Automatic Identification of Experts and Performance Prediction in the Multimodal Math Data Corpus through Analysis of Speech Interaction. In: Second International Workshop on Multimodal Learning Analytics, Sydney, Australia (2013)
6. DMello, S., Picard, R., Graesser, A.: Towards An Affect-Sensitive AutoTutor. IEEE Intelligent Systems 22(4), 53–61 (2007)
7. DMello, S.K., Craig, S.D., Witherspoon, A., McDaniel, B., Graesser, A.: Automatic detection of learners affect from conversational cues. User Model User-Adap. Inter. (2008), doi:10.1007/s11257-007-9037-6
8. Morency, L.P., Oviatt, S., Scherer, S., Weibel, N., Worsley, M.: ICMI 2013 grand challenge workshop on multimodal learning analytics. In: Proceedings of the 15th ACM on International Conference on Multimodal Interaction (ICMI 2013), pp. 373–378 (2013)
9. Qi, F., Bao, C., Liu, Y.: A novel two-step SVM classifier for voiced/unvoiced/silence classification of speech. In: International Symposium on Chinese Spoken Language Processing, pp. 77–80 (2004)
10. Sadegh, M., Ibrahim, R., Othman, Z.A.: Opinion Mining and Sentiment Analysis: A Survey. International Journal of Computers & Technology 2(3) (2012)

11. Saif, H., He, Y., Alani, H.: Semantic Sentiment Analysis of Twitter. In: Cudré-Mauroux, P., et al. (eds.) ISWC 2012, Part I. LNCS, vol. 7649, pp. 508–524. Springer, Heidelberg (2012)
12. Schatten, C., Schmidt-Thieme, L.: Adaptive Content Sequencing without Domain Information. In: Proceedings of the Conference on Computer Supported Education (CSEDU 2014) (2014)
13. Schuller, B., Batliner, A., Steidl, S., Seppi, D.: Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. Speech Communication (2011)
14. Vygotsky, L.L.S.: Mind in society: The development of higher psychological processes. Harvard University Press (1978)
15. Wang, Y., Heffernan, N.: Extending Knowledge Tracing to allow Partial Credit: Using Continuous versus Binary Nodes. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS, vol. 7926, pp. 181–188. Springer, Heidelberg (2013)
16. Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., Picard, R.: Affect-aware tutors: recognising and responding to student affect. Int. J. of Learning Technology 4(3/4), 129–164 (2009)
17. Worsley, M., Blikstein, P.: What's an Expert? Using Learning Analytics to Identify Emergent Markers of Expertise through Automated Speech, Sentiment and Sketch Analysis. In: Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011), pp. 235–240 (2011)

# Technology Enhancing Learning: Past, Present and Future

Ricardo Kawase, Patrick Siehndel, and Ujwal Gadiraju

L3S Research Center, Leibniz University Hannover, Germany
{kawase,siehndel,gadiraju}@L3S.de

**Abstract.** Every year the European Conference on Technology Enhanced Learning (ECTEL) gathers state-of-the-art research in the TEL field. Eight years have passed since the first edition of this conference, resulting in over 500 research papers published and more than 1000 researchers involved. However, bringing together two different fields of study (Technology and Learning), does not necessarily imply interdisciplinary research. To inspect ECTEL's interdisciplinarity and related facts, we dedicate this paper to study the evolution of the conference over time. In this paper, we provide a thorough analysis of the evolution of papers, authors and topics explored over the years. Our analysis provides an understanding of the origin of the conference and the direction that future research in TEL is moving towards. In addition to this, we built interactive online interfaces to enable researches to explore all the information pertaining to past ECTEL research. These interfaces enable users to easily browse through ECTEL papers, authors, knowledge and connections, possibly leveraging the discovery of related work and future collaborations.

## 1 Introduction

Every year since 2006, the European Conference on Technology Enhanced Learning (ECTEL) has been gathering the elite research papers and researchers in the field of *technology* and *learning*. Since its first year, the conference has had a clear goal; to bring together researchers from the fields of technology and learning, and provide a discussion forum to blend both fields [13]. In fact, year after year, the conference is successfully achieving its founders' vision. In eight years of this conference, over 500 research papers have been published with more than 1000 authors involved.

However, bringing together two different fields of study does not necessarily imply interdisciplinary research. Although many papers provide a fine balance between technological and learning advance, browsing past proceedings we can clearly observe papers that are mainly focused on learning aspects, e.g. [7,11], and papers that improve the state-of-the-art technology but are not strictly linked to learning scenarios, e.g. [8,3].

In addition to this observation, during the ECTEL 2013 conference, we informally interviewed several attendees who have confirmed witnessing a gap between learning and technology within the conference. These observations motivated us to explore the past, current and future standing of ECTEL conference.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 193–206, 2014.
© Springer International Publishing Switzerland 2014

Since the main goal of ECTEL is to augment the interdisciplinarity between technology and learning, we believe that, by supporting researchers from both fields in easily finding and exploring past elements of the venue, we can significantly contribute to achieve this goal.

In this paper, we present a thorough analysis of eight years of ECTEL, including general statistics of the conference, distinctive analysis of the learning and technological fields, co-authoring analysis, community evolution, conference topic evolution among other interesting facts.

The metadata were extracted from ECTEL Webpages[1], the DBLP Computer Science Bibliography[2], and the Digital Library of Springer[3]. Further, all papers were downloaded in PDF format, converted to text and semantically annotated using the Wikipedia Miner tool [12].

In addition to the data analysis, we provide first results in the direction of predicting new community members and *hot* topics. Finally, to conclude our work, we provide an online interactive interface where researchers and everyday users can browse contents of ECTEL. The interface provides an easy-to-use interactive overview of the conference statistics, authors' profiles, collaborations network, keywords and annotations. Users can browse through the different pages, finding interesting papers, relevant related work and potential collaborations.

To summarize, our main goal is to provide a settlement between technology and learning, thus, the main contributions of this paper are outlined as follows:

- (i) An analysis of all ECTEL research works published to date.
- (ii) Annotated data of all the content from ECTEL papers.
- (iii) Feature selection for predicting new community members and conference *hot topics*.
- (iv) Online interactive interfaces to support the exploration of ECTEL data.
- (v) Exposure of all free available data to the Linked Open Data cloud for re-use.

The rest of this paper is structured as follows. In Section 2 we discuss past literature related to our work. Section 3 is dedicated to explain the data collection process and the analysis of interesting facts pertaining to past editions of ECTEL. In Section 4, we present experiments done in the direction of predicting future community members and popular topics in the conference. In Section 5, we briefly describe the user interfaces that enable users to browse ECTEL authors, publications, network and knowledge. Finally, in Section 6, we conclude our work and discuss future directions.

## 2 Related Work

Every few years, research and contributions based on bibliometrics aspects emerge in specific communities. For example, in the field of digital libraries, Liu et al. [9] analyse the contents from the Joint Conference on Digital Libraries community. Similar to our

[1] http://www.ec-tel.eu/
[2] http://dblp.uni-trier.de/
[3] http://www.springerlink.com/

work, they expose several interesting facts of the DL community and an additional model to rank influential authors (AuthorRank).

In the field of Human Computer Interaction (HCI), Henry et al. [6] provide an analysis based on 4 major conferences in the field. The paper provides several visualizations that help readers to understand the characteristics of the HCI network, however, it does not provide an in-depth analysis or any interactive interface.

ACM Hypertext conference series also had its own analysis published by Chaomei Chen and Les Carr [1]. There, the authors present citations and co-authorship analysis from nine conferences over ten years. Their findings show that almost half of the papers refer to papers from the same series, which points to a very homogeneous research community.

Coming to the TEL field, Ochoa et al. [14] provided a co-authorship and citation analysis focused on TEL publications presented at EDMedia conferences. Similar research was also performed by Fisichella et al. [4] on top of partial data from EDMedia and ECTEL conferences. Their results provide interesting insights regarding collaboration networks in the TEL area and support the importance of co-author analysis and citation analysis for understanding, as well as analyzing scientific communities.

More recently, active members of ECTEL community have published interesting analysis on TEL bibliometrics. Reinhardt et al. [15] provided a brief authorship, co-authorship and citation analysis based on the first five years of ECTEL. With a slightly different approach, Derntl and Klamma [2] presented a thorough social network analysis on European TEL projects. Although focused on data from projects, their report is extremely relevant for the TEL research community. In fact, as we show in Section 3, TEL projects play a major role in supporting research.

Although plenty of research has been done on analyzing characteristics of conferences, authorship, co-authorship and citations, only a few provide up to date data that can be accessible by researchers. Thus, the main differential of our work is that, we not only provide an up to date analysis of eight years of ECTEL conference, but we also provide usable tools that allow researchers to explore the data, supporting the related work finding and possibly leveraging collaboration. Additionally, we provide enriched annotated data, an individual analysis of *technology* versus *learning*, and expose the data to LOD cloud.

## 3 ECTEL Uncovered

The first step in our work consisted of collecting data regarding past events of ECTEL. The entire data collection process consists of several steps that help us to access different online repositories. In the first step, we build a reusable[4] crawler that downloads pages with the metadata available in the DBLP Computer Science Bibliography, in our case, the ECTEL content pages[5]. The crawler downloads and structures the information such as year, paper title, number of pages, authors and DOI (Digital Object Identifier) url.

[4] Our crawler can be applied to different conference venues.

[5] http://www.informatik.uni-trier.de/~Ley/db/conf/ectel/index.html

**Table 1.** General statistics of ECTEL over the years. (*unique persons)

| Year | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | Total |
|---|---|---|---|---|---|---|---|---|---|
| Papers | 76 | 46 | 52 | 85 | 67 | 50 | 65 | 92 | 533 |
| Authors | 247 | 155 | 171 | 281 | 276 | 214 | 265 | 387 | 1193* |
| Citations (Avg) | 12.1 | 12.8 | 8.3 | 7.5 | 5.5 | 5.0 | 2.6 | 1.4 | 7.8 |
| Committee Members | 25 | 36 | 29 | 56 | 85 | 80 | 73 | 92 | 157* |

In the second step, we crawl each paper's DOI URL hosted at the Digital Library of Springer. Each paper has its own page containing further metadata. From Springer we collected abstracts of the papers, authors keywords, authors' affiliations, and the entire paper in PDF format[6].

The third step consists of extracting the information within the PDFs files. To this end, we converted the files to text format and extracted their contents and the acknowledgements sections.

In the fourth step, we enrich the text of the papers with annotations of entities. To perform this task, we use the WikipediaMiner [12] tool, a Web annotation service that is responsible for identifying all mentions of entities that can be linked to Wikipedia articles. Basically, the WikipediaMiner algorithm consists of two phases. Firstly, it detects and disambiguates words in the text that represent links to Wikipedia. To disambiguate, WikipediaMiner relies on machine learning algorithms that take into consideration the context of the word. Next, based on the first phase, their algorithm creates links from the disambiguated words to Wikipedia articles. Only those words that are considered to be relevant for the entire document are linked to the corresponding articles. The goal of the whole process is to annotate a given document in the same way as a human would link a Wikipedia article.

As a result, we have ECTEL research papers annotated with entities. The goal of this annotation is to provide a relational knowledge base of the real content of each paper. In contrast to authors's given keywords, the annotations can precisely characterize the contents of a paper and the profile of an author.

As a next step, we searched and crawled Google Scholar[7] to find out the number of citations of each paper. Google Scholar provides a good approximation of numbers of citations, however, one should be careful when drawing strong conclusions based solely on these numbers - it has been proved that this information is susceptible to manipulation [10].

In a final additional step, we crawled the Webpages of each ECTEL conference (from the year 2006 until 2013) in order to download the information regarding program committee members. All collected data was organized and stored in a relational database.

The general statistics of the ECTEL conference are exposed in Table 1. Although the last edition of ECTEL (2013) might be considered the most successful in number of papers and authors, we cannot observe a noteworthy growth over the years. However, program committee growth depicts how the community has expanded. We additionally structured the acknowledgment sections of each paper. In total, 48% of all papers have

[6] Note that PDFs are not freely available.

[7] `http://scholar.google.com/`

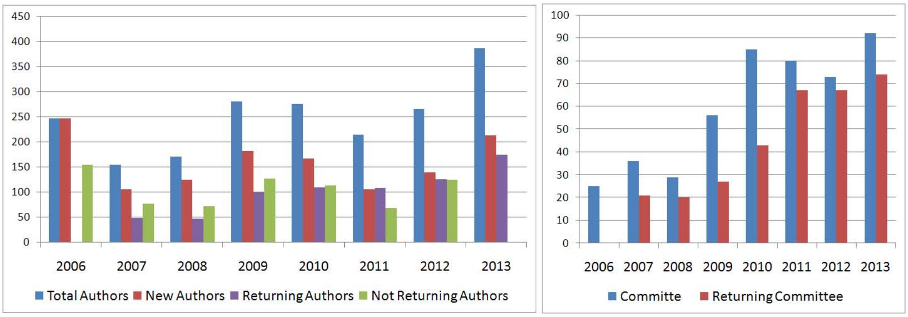



**Fig. 1.** On the left, the number of authors each year of ECTEL. *New Authors* represent authors that never published in ECTEL before. *Returning Authors* represent authors that have at least one publication in a previous year. *Not Returning Authors* represent authors that published only in that give year. On the right, the committee members numbers for each year of ECTEL. *Returning Committee* are the members who have been in part of the committee in a previous year.

an acknowledgment section, 26% of them mention projects support and 20% explicitly mention European Commission support.

### 3.1 Community Growth

The establishment and growth of the ECTEL community is better depicted in Figure 1. We see the increasing number of *Returning Authors* and *Returning Committee*, meaning that each year the community welcomes more and new researchers. However, this growth is in fact smoother than it seems. The distribution of papers and authors follows a power law distribution where 833 authors (69.8%) have only one publication in ECTEL. Out of these 833 authors, 239 (29%) are listed as first authors in their papers.

In Figure 2, we plot controversial yet interesting results. The Figure draws together the number of accepted papers for each year, together with the number of program committee members and the number of papers that had at least one program committee member as author. In average, 49% of program committee members have papers accepted at the conference in the same year (varying from 36% up to 64%). Their papers consist of around 38% of all papers accepted over the years. In fact, the highest numbers were observed in the last edition of ECTEL (2013), where 53% of all papers belonged to at least one of the program committee members.

On one hand, these numbers raise a flag showing that the community is isolating itself. On the other hand, it shows the strength, unity and homogeneity of the TEL research community, aligned with findings from Chaomei Chen and Les Carr [1] on the internal citations on ACM Hypertext conference series. Those who are considered experts in the field (committee members) are constantly contributing to the progress of knowledge on TEL.

### 3.2 Topic Evolution

In order to understand the main themes discussed in ECTEL and the shift in topics over the years, in Figure 3, we see the tag clouds of authors' keywords choice for their

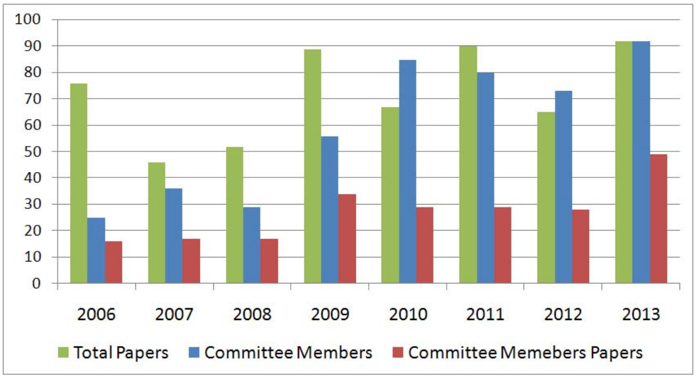



**Fig. 2.** Total number of ECTEL papers, committee members and papers from committee members over the years

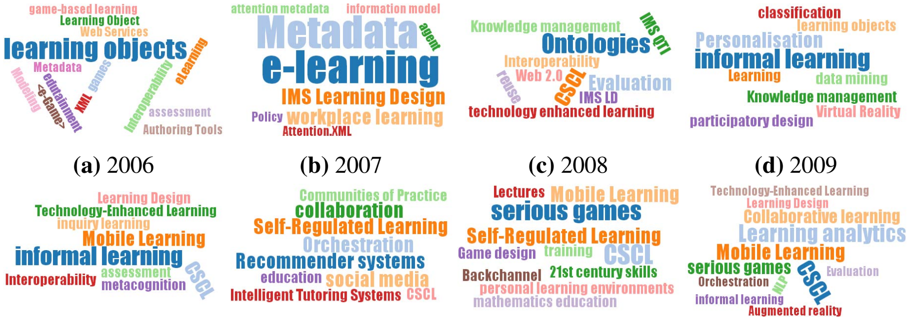
papers. We see that, in the beginning (2006 and 2007), the conference had several publications on learning resources (*learning objects*) and information extraction (*metadata*). Later, in 2008 we see the rise of information management (*ontologies* and *knowledge management*). *Informal learning* was a popular topic during 2009 and 2010 venues, later giving way to *self-regulated learning*. The topic *mobile learning* first appears in 2008 without great impact, however it comes into the spotlight in 2010, maintaining its significance till date.

We also identify the appearance of *social media* in 2011, however, the topic did not catch up in the community. On the other hand, *serious games* established itself as a popular topic in the past two years. Finally, we see that *learning analytics* became, the most discussed theme in the last edition of ECTEL out of the blue. We believe that this was caused by the emergence of another venue in 2011, the International Conference on Learning Analytics and Knowledge (LAK). This venue is mainly organized by the same key persons of ECTEL and it is a natural trend that popular topics of discussion might migrate from one venue to another. However, the analysis of LAK conference is out of the scope of this paper.

### 3.3 Technology versus Learning

In this paper, we are also interested in understanding the individual evolution of the fields of technology and learning, and how they have blended together over the years. To this end, we manually annotated 1,197 keywords assigning a value to each one, whether it is a technology related term, a learning related term, or a neutral one. Each term was evaluated by three annotators, and the final annotation was decided by majority voting. In case of conflict where each annotator assigned a different value, the term was considered neutral.

For example, terms such as *metadata*, *ontologies*, *web services*, *semantic web*, among others, have clear technological characteristics. On the other hand, terms such as *Orchestration*, *autism*, *reflective learning*, *phenomenography*, etc., provide a stronger association with the *learning* field. A few examples of keywords classified as *neutral* are *infrastructure*, *feedback*, *survey*, *TEL*.




**(a)** 2006 **(b)** 2007 **(c)** 2008 **(d)** 2009

**(e)** 2010 **(f)** 2011 **(g)** 2012 **(h)** 2013

**Fig. 3.** Tag clouds of papers' keywords (given by the authors) over the years

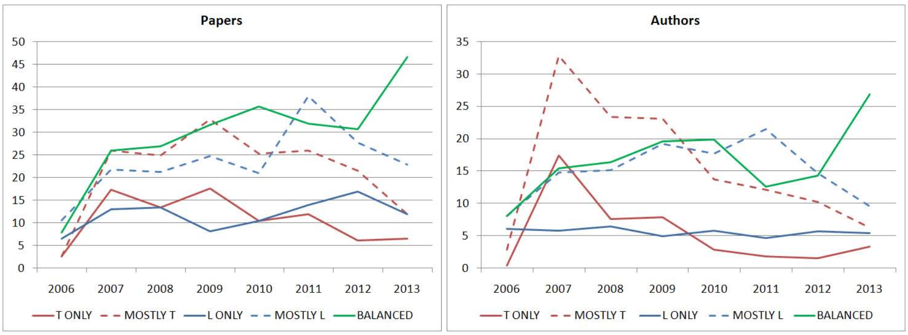



**Fig. 4.** The percentage of papers (on the left) and authors (on the right) pertaining to the fields of mostly or exclusively to the fields technology or learning, and balanced ones. *T ONLY* means papers/authors that have only technological keywords associated. *MOSTLY T* represents papers/authors that have 2 or more technological keywords than learning keywords. *L ONLY* means papers/authors that have only learning related keywords associated. *MOSTLY L* represents papers/authors that have 2 or more learning related keywords than technological keywords. Finally, *BALANCED* are papers/authors that have the approximately same amount (equal or ±1) number of technological and learning related keywords associated.

We found out that each paper has in average 4.6 given keywords. Figure 4 shows the evolution of each field (technology versus learning) over the years in ECTEL. In terms of papers, we see that *exclusive* and *mostly learning* related research had significant growth in 2011 and in 2012. On the other hand, authors who were *exclusively* or *mostly* on the field of technology are slowly disappearing, or integrating their knowledge into the educational field. The rather constant increase of *BALANCED* papers and authors (those who have a fine equilibrium between technological and learning related keywords) depicts the accomplishment of ECTEL in bringing together both fields. The number of papers and authors that unify both fields has never been as high as in the last edition (2013) of ECTEL.


**Fig. 5.** Tag cloud of the annotations

### 3.4 Annotations

The annotation process with Wikipedia Miner resulted in a total of 171,281 annotations. We identified 24,004 unique terms with an average of 321 terms per research paper. The tag cloud in Figure 5 depicts the top discussed topics in ECTEL research papers. The main difference between this tag cloud and the ones presented in Figure 3 is that the annotations are based on the whole content of the papers. It better represents what lies inside each research paper, rather than keywords chosen by the authors. Based on these annotations we are able to build a connected knowledge graph which will allow us to browse and explore the knowledge provided by papers and authors (see Section 5). Additionally, this data will be used in Section 4 for the prediction tasks.

## 4 Forecasting ECTEL

In this work, we are also interested in discovering whether it is possible to predict future characteristics of ECTEL conferences based on historical data. To tackle this, we chose two distinct experimental tasks: (i) predicting new community members and (ii) predicting future conference hot topics. Among the many possible predicting tasks, we consider these to be of greater interest for the community. The experiments presented in the remainder of this section were performed using the machine learning framework WEKA [5] and its implementation of Naive Bayes classifier.

### 4.1 New Community Members

For the first task, we are interested in investigating if a given author will, at some point in the future, *return* to the conference, i.e. will have another publication in a future edition of the ECTEL. Although debatable, we assume that a person who has published more than once (in different years) at ECTEL, has joined the community. In this light, we name this task *predicting new community members*.

For each author's first appearance, we arranged the data in order to aggregate past evidence that might help us predict his/her return. The data consists of pin-point information of the authors' first appearance, e.g. type of paper (short/full), authors' order, keywords, annotations, presence in the program committee and co-authors. Although we are considering the first appearance of a given author, his/her co-authors might have participated in ECTEL in previous editions. Therefore, to a given user profile we also consider if his/her co-authors had papers before, how many and if they were members of the program committee.

With this set of features, we were able to predict new community members with an accuracy of 81.9% (precision=0.826, recall=0.819 and f-measure=0.79). Surprisingly, the only features that have a significant weight in the decision process are *paper_type* and *was_pc_member*. These results mean that an author of a full paper has a higher probability of returning to the conference in a following year. The same applies for an author that is member of the program committee. Besides these features, the experiments detected several co-authors that have worked with returning authors. Examples of the top classified co-authors are Yannis A. Dimitriadis, Gonzalo Parra and Erica Melis.

### 4.2 Predicting Hot Topics

Our second chosen task consists of discovering which historical data are the most prominent for predicting future *hot topics* (papers' keywords chosen by the authors). For example, as we have seen in Figure 3, *personalization* and *informal learning* were hot topics in the 2009. Lately, in the year 2013, *learning analytics* and *mobile learning* became hot topics. Thus, the question raised is if it is possible to identify which past evidence leads to the rise of such topics. For example, if particular authors influence research trends.

For each appearance of a keyword in a particular year, we aggregate the information of this keyword pertained to the three previous years. Given that *learning analytics* is a hot topics in 2013, we investigate the information related to this topic from the years 2010, 2011 and 2012.

In this experiment, we were able to predict *hot topics* with an accuracy of 91.8% (precision=0.974, recall=0.918 and f-measure=0.941). Table 2 depicts the top influential features. We found out that the number of authors that chose a given topic two years in the past is the topmost influential feature, followed by the number of co-occurring topics in the previous two years. It is interesting to see that features regarding two years in the past have a higher impact than features extracted from a single year before. This means that topics that emerge in a given year, take another two years to become a *hot topic*.

Additionally, we observed that facts regarding a given topic that are older than two years do not provide useful information for predicting a *hot topic*. Finally, the experiments also output several authors who are ahead in time. These authors have published works of a given topic before it became popular. Marcus Specht, Nicole C. Krämer, Davinia Hernández Leo and Günter Beham amongst others might be considered visionaries in the ECTEL community.

**Table 2.** Top features for predicting new ECTEL hot topics

| Rank | Feature Name (years before) | Weight |
|---|---|---|
| 1 | # of authors (-2) | 0.04077 |
| 2 | # co-occurent Keywords (-1) | 0.03652 |
| 3 | # co-occurent Keywords (-2) | 0.03368 |
| 4 | # of papers (-2) | 0.03324 |
| 5 | # of authors (-1) | 0.02721 |
| 6 | # of papers (-1) | 0.02721 |

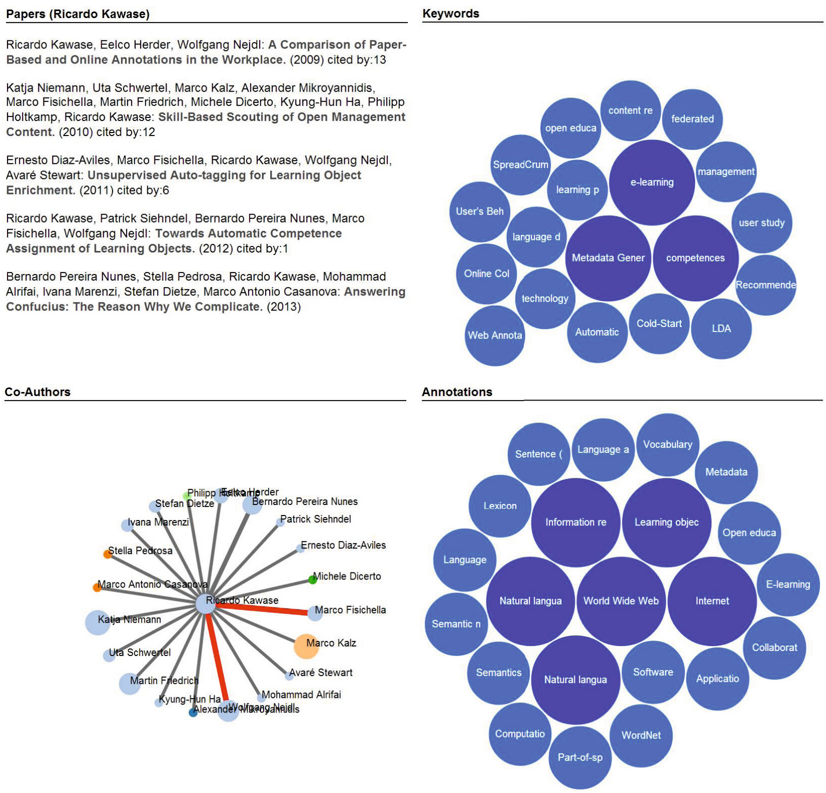



**Fig. 6.** Authors profile interface containing list of papers, keywords, annotations and collaborations

## 5 Exploring ECTEL

In order to expose our results and analysis to the community, we built several online interactive interfaces that enable users to browse the contents of ECTEL through different facets views.

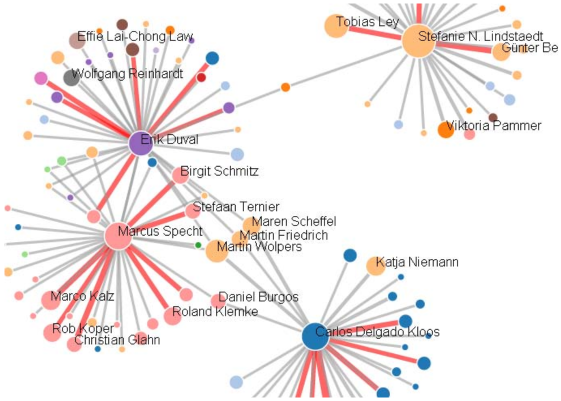



**Fig. 7.** Iteractive collaboration graph

We built all interactive graphic visualizations using D3 (Data-Driven Documents) JavaScript library[8]. The visualizations we present consist of bar charts, tag clouds, force-directed graph, and bubble charts among others. The interfaces provide information that is easy to digest regarding statistics of the conference, content of the papers and relationships between authors, papers and knowledge.

It provides easy access and navigation to all sections of the website, an easy to use and attractive interface and supportive help texts. Additionally, it was developed to be platform independent and to support any modern browser in any device.

While it is not hard to find relevant research using existing tools, there is little one can find beyond explicitly specified authors and corresponding publications. The main contribution of our interfaces is the assistance to users, for discovering information, authors' network connections and relevant related work.

The Website is divided in four main categories: Authors, Collaboration, Keywords and Knowledge. *Authors* section provides the profile of each author in the ECTEL community. It lists information of the authors papers, the keywords usually associated to the author, annotations and the collaboration graph (see Figure 6). Additionally, it provides an overview of the author publication count in comparison to the rest of the community.

In the *Collaborations* section, one is able to clearly identify all co-authoring relations of a given researcher, or simply browse through the whole co-authoring graph (see Figure 7).

In the *Keywords* section, the user can identify which are the most trending topics of a conference and the most influential researchers in those topics. The interface further helps to track evolution in the research fields, by leveraging year-based filters to help users in identifying the transition and development of knowledge and authors along the years (see Figure 8). Both authors' keywords and annotations are provided in this section.

The section *Knowledge* provides similar interfaces as the section *Collaboration*, however, the connections are based on keywords/annotation co-occurrence. This section provides great exploratory features for finding relevant related work and experts on

[8] `http://d3js.org/`

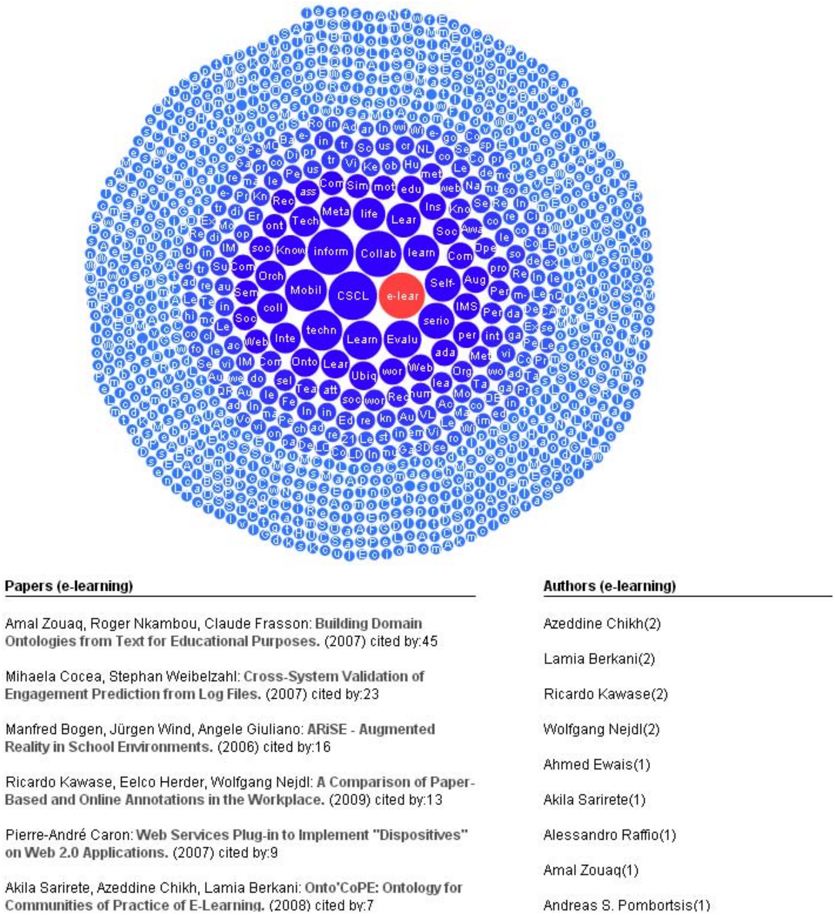



**Fig. 8.** Iteractive keywords bubbles interface

each field of knowledge. All the interfaces provide interactive commands that trigger browsing, exploration and filters.

We invite the reader to try our online prototype[9].

### 5.1 ECTEL in the LOD Cloud

By bringing the underlying complex relations (between authors, research collaborations and conferences) to the surface, we greatly improve the user experience while immediately satisfying one's information need. As an additional contribution, we expose this enriched knowledge as Linked Data (by following the principles of publishing such

[9] http://www.l3s.de/~kawase/DBLPXplorer/ECTEL/

data) that can be queried at the SPARQL endpoint[10]. We also provide for a Pubby[11] interface to facilitate additional exploration of the knowledge base. The interface can be accessed at http://meco.l3s.uni-hannover.de:8899/dblpXplorer/. By exposing this knowledge to the Linked Open Data Cloud, we promote re-use of the structured data.

## 6 Conclusion

In this paper, we presented a thorough data analysis on top of the past eight years of ECTEL. First, we described, step by step, the data crawling and enrichment process. It is important to emphasize that your crawling steps presented in Section 3 are easily reusable, thus enabling us to build similar knowledge over difference conference venues.

Based on the collected data, we uncovered interesting facts of the TEL community regarding community growth, topics evolution in the conference and a breakdown comparison between technology and learning. Additionally, we performed two experiments in order to demonstrate that the historical data of ECTEL can be used for predicting upcoming facts with relatively high accuracy.

In fact, our main contribution lies in the exposure of the collected data back to the community. Our interactive user interfaces allow researchers and the general interested public to browse through the available and enriched data past venues of ECTEL. Our user interfaces leverage the finding of related work, experts on different topics, and most importantly, might augment collaborations. Additionally, the exposure of the collected data to the linked open data cloud allows third-part applications to reuse our work.

Although we identify our work as a contribution to the community, it indirectly contributes to improvement of research in the TEL field. As future work, we plan to conduct a user evaluation of our proposed interfaces in order to collect feedback. Based on the given feedback we plan to improve the interfaces and to build new means for browsing the data.

**Acknowledgement.** This work was partly funded by the LinkedUp project (GA No: 317620) under the FP7 programme of the European Commission.

## References

1. Chen, C., Carr, L.: Trailblazing the literature of hypertext: Author co-citation analysis (1989–1998). In: Proceedings of the Tenth ACM Conference on Hypertext and Hypermedia: Returning to Our Diverse Roots: Returning to Our Diverse Roots, HYPERTEXT 1999, pp. 51–60. ACM, New York (1999)
2. Derntl, M., Klamma, R.: The european tel projects community from a social network analysis perspective. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 51–64. Springer, Heidelberg (2012)

[10] http://meco.l3s.uni-hannover.de:8829/sparql?default-graph-uri=http://purl.org/dblpXplorer/

[11] http://wifo5-03.informatik.uni-mannheim.de/pubby/

3. Diaz-Aviles, E., Fisichella, M., Kawase, R., Nejdl, W., Stewart, A.: Unsupervised auto-tagging for learning object enrichment. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 83–96. Springer, Heidelberg (2011)
4. Fisichella, M., Herder, E., Marenzi, I., Nejdl, W.: Who are you working with? - visualizing tel research communities -. In: Proc. of World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2010), Toronto, Canada, June 28-July 2 (2010)
5. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1), 10–18 (2009)
6. Henry, N., Goodell, H., Elmqvist, N., Fekete, J.-D.: 20 years of four hci conferences: A visual exploration. Int. J. Hum. Comput. Interaction 23(3), 239–285 (2007)
7. Krogstie, B.R., Prilla, M., Pammer, V.: Understanding and supporting reflective learning processes in the workplace: The CSRL model. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 151–164. Springer, Heidelberg (2013)
8. Lemnitzer, L., Vertan, C., Killing, A., Simov, K.I., Evans, D., Cristea, D., Monachesi, P.: In: EC-TEL, pp. 202–216 (2007)
9. Liu, X., Bollen, J., Nelson, M.L., de Sompel, H.V.: Co-authorship networks in the digital library research community. Information Processing & Management 41(6), 1462–1480 (2005); Special Issue on Infometrics
10. López-Cózar, E.D., Robinson-Garcia, N., Torres-Salinas, D.: Manipulating google scholar citations and google scholar metrics: simple, easy and tempting. CoRR, abs/1212.0638 (2012)
11. McLaren, B.M., et al.: To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 222–235. Springer, Heidelberg (2012)
12. Milne, D., Witten, I.H.: Learning to link with wikipedia. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, pp. 509–518. ACM, New York (2008)
13. Nejdl, W., Tochtermann, K. (eds.): EC-TEL 2006. LNCS, vol. 4227. Springer, Heidelberg (2006)
14. Ochoa, X., Méndez, G., Duval, E.: Who we are: Analysis of 10 years of the ed-media conference. In: Siemens, G., Fulford, C. (eds.) Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009, pp. 189–200 (2009)
15. Reinhardt, W., Meier, C., Drachsler, H., Sloep, P.: Analyzing 5 years of ec-tel proceedings. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 531–536. Springer, Heidelberg (2011)

# Immersive Multi-user Decision Training Games with ARLearn

Roland Klemke, Stefaan Ternier, Marco Kalz, Birgit Schmitz, and Marcus Specht

Welten Institute – Research Center for Learning, Teaching and Technology,
Open University of the Netherlands, P.O. Box 2960, 6401 DL Heerlen, The Netherlands
{roland.klemke,stefaan.ternier,marco.kalz,
birgit.schmitz,marcus.specht}@ou.nl

**Abstract.** Serious gaming approaches so far focus mainly on skill development, motivational aspects or providing immersive learning situations. Little work has been reported to foster awareness and decision competencies in complex decision situations involving incomplete information and multiple stakeholders. We address this issue exploring the technical requirements and possibilities to design games for such situations in three case studies: a hostage taking situation, a multi-stakeholder logistics case, and a health-care related emergency case. To implement the games, we use a multi-user enabled mobile game development platform (ARLearn). We describe the underlying real world situations and educational challenges and analyse how these are reflected in the ARLearn games realized. Based on these cases we propose a way to increase the immersiveness of mobile learning games.

**Keywords:** mobile learning, game-based learning, multi-user games, decision processes, multi-role game-design, immersiveness.

## 1 Introduction

Due to their ability to teach innovatively and to deliver needed skills effectively, serious games received a high interest in recent years [1]. Recent investigations into educational games focus on the motivational potential and their low-threshold learning opportunities [2][3] as well as their ability to address various target groups [4][5]. The Mobile Learning NETwork's (MoLeNET) review on learning game technologies suggests that mobile learning games provide potential for learning and teaching in terms of 'assessment', 'learner performance and skills development' or 'social and emotional well-being' [6]. Even though multi-user gaming environments are around for many years, little research has been shown to exploit multi-user enabled platforms systematically for learning. Approaches available so far mainly concentrate on the use of virtual worlds for multi-user games [7][8]. Multi-stakeholder decision situations confronted with time restrictions and incomplete information such as emergencies have been recognised as a relevant field for specific training approaches involving tabletop exercises [9] or (non-computerized) tactical decision training games [10].

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 207–220, 2014.
© Springer International Publishing Switzerland 2014

First prototypes towards the use of collaborative computer games are also reported [11][12]. The systematic use of game platforms is still in its infancies.

In this paper, we explore the use of a multi-user enabled mobile serious gaming platform to a number of immersive, real-world, multi-stakeholder decision training situations. We describe three real world situations. For each, we analyse the educational challenges connected to the training of these situations and describe the game-design we used to address these challenges.

The focus of this paper is to explore common requirements for the training of very different decision situations and showing their implementation in real-world cases. Consequently, we ommit some individual details of the cases, which have been published separately for each case (see references below), including first user experiences. The three decision cases we use are:

1. Cardiac arrest emergency involving bystanders [13].
2. Stakeholders in a hostage taking case of a distributed organization [14].
3. Disturbances in a complex logistics value chain [15].

In the following section we will briefly introduce and motivate the three cases and the challenges involved. We will subsequently reflect on related work and derive our requirements to support the described educational scenarios. We introduce the technology used in our cases followed by the game-design applied. Finally, we compare the three cases and discuss our findings in the light of upcoming technologies.

## 2 Background

In this section, we highlight background information about the three real world cases introduced above and motivate the game-based learning approach.

### 2.1 Bystander Decisions Processes in Emergency Situations

Cardiac arrest is one of the main causes of death worldwide. The rate of first-responder CPR is critical to increase survival rates since the professional medical emergency services need approximately 8 – 10 minutes to arrive at an incident. The project EMuRgency aims to increase the rate of bystander resuscitation and thus survival chances by socio-technical innovations. One of these innovations is the use of a training game. Traditional training approaches for pre-hospital resuscitation training are mostly based on lecture-centric phases in combination with training of motor-skills on a mannequin in a group-training context. In many European countries this training is part of a mandatory first-aid-training before being allowed to get the driver license. While this is on the one hand a window of opportunity to train large parts of a population in basic resuscitation skills and knowledge from an educational perspective this training format delivers only short-term knowledge and competence building whose retention times is normally not longer than 3 to 6 months [13].

## 2.2 Hostage Taking Situation

The Office of the United Nations High Commissioner for Refugees (UNHCR) leads and co-ordinates international action to protect refugees and resolve refugee problems worldwide. UNHCR staff often operates in hazardous locations. As this organisation is consequently confronted with kidnappings of their co-workers, employees are trained on how to deal with such situations. The Global Learning Centre (GLC) of the UNHCR based in Budapest organizes these trainings. The typical approach is a workshop organized over a 3-5 day period in which different aspects of security risk management are addressed. These workshops cover policy-based information, such as standard operating procedures, delivered through self-study pre-E-learning modules followed by instructor-led workshops. The workshop also includes immersive simulation exercises, for example hostage taking, bomb threat and other security-related scenarios.

While for many years a role-playing game has been part of these workshops, running the game turns out to be an intensive exercise, not only for the participants, but also for the organizers and facilitators. As the game is carried out at a rapid pace it can be difficult to have an all-inclusive debriefing in which all roles of all teams receive appropriate feedback. This aspect led to the development of this project, in an effort to address this shortcoming. A technology-based alternative for the original game should address the following training issues [14]:

- Enabling the creation of different reusable variations of a game-design for emergency security response, covering initially the hostage taking situation and potentially other cases.
- Enabling ‘on the fly’ messaging to participants and real-time assessments of activities.
- Semi-automatic management of the game thus enabling more participants to experience the role-playing exercise.
- Creating a log through the game of responses and interactions, which can be used by the trainer to provide feedback during the debriefing session.

## 2.3 Decisions in Logistical Value Chain

In a huge international port, like the Port of Rotterdam, thousands of containers are moved every day in and out through several different channels. Different interests have to be met during these operations, which are reflected by various stakeholders, who are equipped with different responsibilities and partly conflicting goals and who have to interoperate: the *control tower* needs to ensure the overall smooth operation, the *resource planner* is responsible for organizing and assigning the internal port personnel, the *yard planner* is responsible for the storage of containers in the port’s internal storage places, the *vessel planner* is responsible to deliver container to and from vessels, the *sales manager* is interested in customer satisfaction.

Disturbances (such as delays, malfunctioning machinery, accidents, strikes) may cause severe ripple effects resulting in high costs. The situation gets worse, when the operating individuals are not aware of interdependencies and conflicting goals.

The SALOMO project aims to provide a game-based training solution to create shared situational awareness [15] to cope with this situation and to highlight the importance of communication.

## 3 Related Work

Decision-making in sociotechnical systems (large technical systems involving many stakeholders) is complex and error-prone due to inter-dependencies and lack of information [16]. Additional situational information might help to gain shared situational awareness (i.e. "a common relevant picture distributed rapidly about a problem situation" [15]). Therefore it is crucial to understand the role of communication and inter-dependencies among stakeholders [17].

Several educational theories are related to the goal of embedding learning processes into real world application and performance. The anchored instruction approach [18] was developed to decrease the problem of inert knowledge through the presentation of real authentic problems and the active exploration by learners. The theory of situated learning [19] 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. Learning games provide such environments, in which learning processes can be embedded in situations similar to real life. They provide realistic problem situations and allow players to actively explore solution paths. Multi-user games can also provide the social context, in which learning takes place.

Learning games provide an environment, in which learning processes are embedded in situations similar to real life: in her review of immersive games, [20] stresses the importance of linking the experiences made in a game, simulation or micro world with their application in real world practices. Game-based approaches towards the distribution of knowledge for emergency situations can also be found. However, many of them focus on the factual knowledge rather than the decision process [21]. In the approach presented here, we aim to combine factual knowledge provision with decision training.

The importance of specific training towards fast and process decisions in emergency situations has been addressed by approaches such as tabletop exercises [9]. To improve decision training, specific training games have been proposed [10][22][23], in order to put trainees in realistic situations. However, these games are often not computerized and thus lack some of the opportunities computer games offer (such as autonomous playability, tracking of user decisions and actions, scalability). Also, these games often require a human game master to track the game progress [24]. Computerized decision training approaches involve immersive virtual reality scenarios [25][26], which put the player into a realistically modelled situation or agent-based approaches, which aim to model co-player behaviour [27]. While these developments deliver single user games, first prototypes have been successfully created towards decision training using collaborative games [11] and multi-user mixed-reality scenarios [12]. While at first sight the latter looks similar to the approach reported here, it does not rely on a re-usable platform for serious games.

## 4 Requirements

All three cases described share some common attributes. They all involve several persons in the decisions, in all three cases decisions need to be taken quickly and in all cases decisions have to be taken in a situation of incomplete, misleading or wrong information. The different persons involved in the decisions act in different roles, which have or require different information. Success can only be gained when the different persons involved cooperate. To provide a training environment for these scenarios, we derived a number of requirements, which are summarized in this section. We are aware, that these requirements not completely describe the necessary features of a game platform (such as user interface aspects, interactivity elements, game patterns, media support). However, here we focus on requirements relevant for multi-user games and decision training. Also, we omit here details of our user-centred requirements analysis processes for our cases. More detail on the technical requirements can be found in [28], [29], and [30]. Details on our user-centred approach and the involvement of stakeholders are published in [14], [15], and [31].

- (R1) An environment supporting these scenarios needs to be multi-user enabled to support the different participants in the educational process. Multiple users need to be able to play games together using different devices. Users need to have personal views on the game state. Teams shall be supported.
- (R2) Different roles for different participants need to be supported, individualising information visibility, tasks, communication, and process steps.
- (R3) Individual information supply and messages depending on player roles shall be possible. Together with (R2), this allows for personalized games according to player roles. Games can be organised such that only collaboration leads to success.
- (R4) The game process shall allow interweaving player decisions with game events and shall allow semi-automatic game execution. Players shall be confronted with the consequences of their decisions. The game processes designed with the platform need to define alternative paths and decision points.
- (R5) The game process should be supported on mobile devices. Events, notifications, decisions should use standard channels. This requirement supports the immersive character by staying close to the environment used to the players.
- (R6) Re-use of games including variations and simple modifications shall be possible. While not related to gaming, this supports evolutionary game designs.
- (R7) The environment should log game activities for later game reviews, debriefings, and the necessary reflection: the game process can be analysed and decisions taken can be discussed.
- (R8) The environment should offer an immersive game play, which puts the player into a realistic scenario.

Confronted with these requirements, we decided not to design our games within a virtual world but rather aim to provide the games in the real environment.

## 5 Technology

We are designing the training games using the ARLearn-platform. ARLearn is a platform for the design of mobile process-based learning games [28]. The platform consists of an authoring interface that enables game-designers to bind a number of content items and task structures to locations, events, and roles and to use game-logic and dependencies to initiate further tasks and activities. The platform has been recently used for several similar pilot studies in the cultural heritage domain [29]. The cloud-based, Google App Engine hosted ARLearn service is an open-source project that permits others to reuse and contribute.

Various kinds of clients connect to this game engine. The Android client allows for game play in the real world, while the StreetView-based client (called StreetLearn) offers the same game logic a virtual environment [30]. The open architecture of ARLearn also allows for new clients to be developed for future applications. An ARLearn game is a reusable game logic description, comparable to CSCL scripts, which model collaborative learning processes [32][33]. However the ARLearn processes explicitly include game patterns [34] such as competition, collaboration, or scoring into the process design and thus embed the collaborative learning experience in the game context. A game run corresponding to a game defines users grouped in teams. While playing, users generate actions (e.g., “read message”, “answered question”) and responses. This output is also managed within the realm of a run.

Looking at other approaches for mobile serious games, we find a few related approaches. The ARIS platform [35] offers the possibility to author location-based mobile games. While ARIS has been successfully used in several application examples [36], it does not support multi-player/multi-role games. QuestInSitu is a mobile learning platform including authoring which mainly focuses on assessment [37] by putting them into location-based contexts. [38] describes an implementation of a team-enabled mobile gaming platform. The location-based task model allows for linear games, where a new task description follows the previous one. In summary, the following motivation guided the decision to use the ARLearn platform to design the games for the abovementioned scenarios and the derived requirements.

- The ARLearn platform is multi-user enabled and supports multiple stakeholder roles and teams within one game (R1 & R2).
- The game process can be individualised according to player roles, so that incomplete or individual information supply is possible (R3)
- The event-based game model of ARLearn allows designing realistic game processes, which simulate mission critical real-life situations and conditions, placed in an augmented reality scenario (R4).
- Commonly used smartphones (Android, iOS) can be used to play ARLearn games, which simplifies game distribution (R5).
- The authoring interface allows copying and modifying games, making the creation of variations easy (R6).
- ARLearn records user activities and allows reviewing game runs at a later stage (R7).
- The ARLearn platform is location-aware, which allows for realistic game-play settings (R8).

ARLearn supports team play and allows for mixtures of competitive and collaborative games. In the next section we discuss the multi-role-based game-designs, which we have implemented with ARLearn to realize the abovementioned scenarios.

## 6 Game-Design

In this section, we present the game designs underlying the three cases. In all cases the games are organized in a three-phase setup, including an introduction phase, a game-phase and a debriefing phase.

a) *Introduction Phase.* The introduction phase includes technical setup, explanation of game mechanics, content, rules, and aim. Groups are formed, roles assigned and the game is started.
b) *Game Phase.* In the game phase, the teams play the game. Each of the three cases follows a different game process as described in the following sections.
c) *Debriefing Phase.* In the debriefing phase game results, team behaviour, and expected outcomes are reviewed. The logging functionality of ARLearn allows analysing individual behaviour and team performance.

As this paper doesn't focus on introduction and debriefing the following sections describe the game phase of the three cases included in this research.

### 6.1 Heart Run Game for Bystanders in Emergency Situations

The main goal of the heart run game is the acquisition of skills and abilities 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 game-design is oriented on the design recommendations for situated learning scenarios. The tasks involved in the game are aiming to produce a more authentic context for learners than the typical classroom lecture. The game can be played with 2 or 3 players and there are 3 different roles foreseen: A CPR player, a player who documents the performance with video recording and an optional player who is responsible to find and get an Automated External Defibrillator (AED) to the victim.

The game is initiated with a notification-message that informs the CPR player that a victim is in the direct surrounding of the team. The CPR player starts to identify the location of the victim accompanied by the documentation player. During the routing-phase the stress level of the player can optionally be increased with sounds or visuals that represent the decrease of oxygen in the body of the victim. After identifying the victim, the CPR player has to perform the steps required in case of a witnessed cardiac arrest, namely securing of the area, calling for help, controlling the breath and finally starting cardiopulmonary resuscitation (CPR). The documentation player records this process. As an option, the AED player receives the location of a nearby AED and has to find the device and bring it to the location of the victim. Here the players have to scan a barcode to communicate to the system that the AED has arrived at the victim. Now the CPR player and the AED player have to coordinate their action in terms of continuing CPR and at the same time preparing the application of the AED. The documentation player is responsible for recording the performance in the best quality possible. The game is over after approximately 8 – 10 minutes after which the emergency medical services arrive. After this the debriefing is organized in two steps: First the players conduct a self-assessment based on a comparison between their recording and a gold-standard video. Then the players discuss the results of the self-assessment with an available tutor. While we have presented here only one round of the game the players can change roles and play the game again.

More detail on the heart run game-design can be found in [31]. Fig. 1 displays screenshots of the HeartRun game implemented in ARLearn. The information displayed on the screen depends on game state and player situation: only when the player is at the right location or takes the right decision, the corresponding instructions are shown.

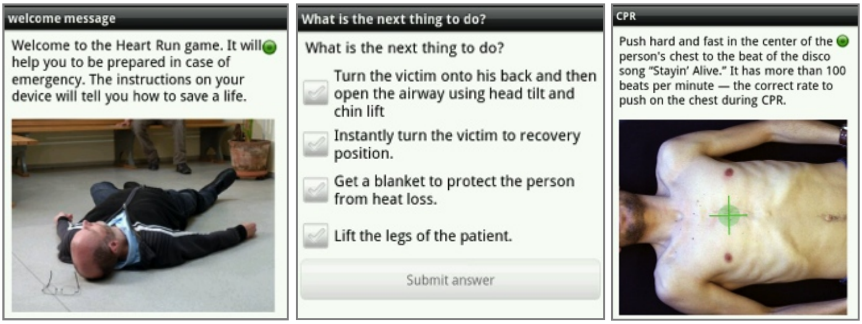



**Fig. 1.** Screenshots of the HeartRun game: welcome message, decision point, and instructions

## 6.2 Hostage Taking Game

The hostage taking game is designed to prepare the participants on the response procedures to be initiated immediately when a staff member is taken hostage. A Hostage Incident Management (HIM) team is deployed eventually in such situations but it can take time till this team arrives and offices need to know how to respond prior to their arrival.

The players participated in this game taking one of the following three roles: head of office, security officer and staff welfare member. The hostage-taking simulation was designed such that players in all roles play the same game but have to react differently based on their roles. The game is organized in 5 rounds.

*Round 1: Notification of the incident.* The game starts with a plea for help by Jerry Khan, a fictitious UNHCR employee that was taken hostage. This video message features a blindfolded actor and creates an authentic context. This message is broadcasted to all the roles. Next, players take a decision on what to do next, depending on their specific role. The head of office (role A) for instance can decide to "notify the Designated Officer (DO)" while a staff-welfare member (role C) should select the option to "contact senior management". Depending on the decision taken, they receive feedback on whether this is a good choice.

*Round 2: Assembling the team.* In the next round, the head of office is informed by the DO that a hostage incident management team will be dispatched. In the mean time, they need to contact the security advisor (role B) and staff welfare officer (role C) and ask them to assemble in headquarters for a planning session.

*Round 3: Planning.* When the facilitator observes that the team has assembled, an audio recording of the DO requesting the team to work out a reception plan is sent out. The team is next tasked to work out this plan on a flip-board and to capture a photo of the plan with their device and submit this as soon as they are ready. Next, the participants are asked to split up and go to their individual rooms.

*Round 4: Responding.* In this round, role A and role C participants are to respond to calls from a journalist and a distressed family member respectively. The security officer (role B) in the meantime receives a message from the DO with the task to prepare a Proof of Life (POL) question.

*Round 5: Negotiating.* In this last round, all roles gather together again. This is triggered by a message from the hostage takers. In this phase, a negotiation with the hostage takers is simulated. The game ends with the message that the Hostage Incident Management (HIM) team has arrived and is ready to take over the negotiations.

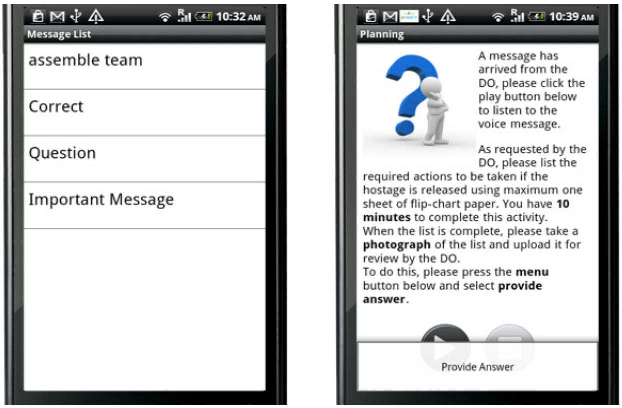



**Fig. 2.** Screenshots of the hostage game: message overview and task description

Fig. 2 shows screenshots of the hostage taking game with an overview of messages and a detail view of a task description and an audio.

### 6.3 Decision Game for Logistical Value Chains

We designed a learning game using the ARLearn-platform. The game is playable with five players in pre-defined roles. Per round, players are confronted with the description of a disturbance situation, which affects the functionality of the port, e.g. a trucker strike. Each player receives a different situation description, depending on the role assigned (while e.g. the resource planner knows about the strike, the yard planner only receives information about missing personnel). Consequently the players need to take decisions according to the individual but incomplete information. Each decision may affect the decisions of other players, which is modelled in terms of score impact. When all players decided in the current round, the game progresses to the next round.

Each level consists of five rounds, which are synchronized after each decision. Each round gives access to a new situation description individualized for each role. While level one of the game isolates the different players completely, subsequent levels give access to limited communicative resources. This shall foster the players to exchange decision related information in order to create awareness for other player's situation and the overall consequences of own decisions.

To provide a realistic and immersive scenario, players can potentially play the game in separate locations as their mobile devices are synchronised automatically via ARLearn. The mobile devices also allow the players to use communication means similar to their daily activities.

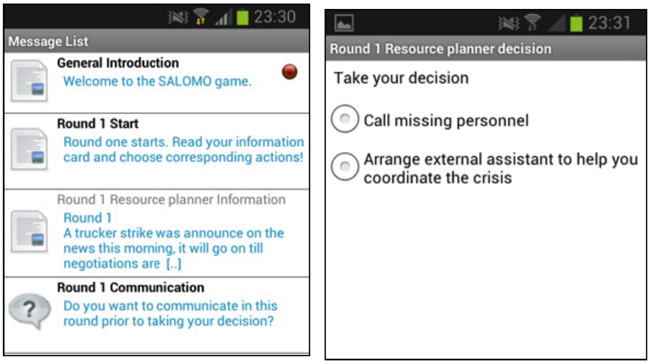



**Fig. 3.** Screenshots of the SALOMO game: message overview and decision point

Fig. 3 displays screenshots of the SALOMO game showing communication messages and decision points.

## 7 Comparison

In this section we compare the three game-designs according to the coverage of requirements listed in section 4 and according to the role-based game design elements used. Table 1 summarizes the coverage of the requirements by the three different cases. All cases use a multi-player, multi-role game-design (R1, R2). Heart Run and the logistics game use the concept of individual information supply (R3), while all games use the concept of event-based notifications (R4). The game is played on standard mobile devices in all three cases (R5). Variations of the game-design are used in the hostage game as well as the logistics game (R6). The logging feature is used in all cases to support the debriefing phase (R7). As opposed to the other two scenarios, in HeartRun the game play does not provide immersiveness (R8): especially, when training goals involve manual operations, players have to switch between device interaction and real world activities too often and perceived these necessary switches as disruption in the game play.

**Table 1.** Mapping of requirements to the game designs

| | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 |
|---|---|---|---|---|---|---|---|---|
| Heart Run | x | x | x | x | x | | x | |
| Hostage | x | x | | x | x | x | x | x |
| Logistics | x | x | x | x | x | x | x | x |

## 8 Increasing the Immersiveness: Discussion and Future Work

From three different real world cases, which cope with complex decision situations involving multiple stakeholders, we have derived requirements to model such situations in a multi-user, multi-role mobile game environment. We have designed multi-user games for these three cases and realized them in the ARLearn platform, which covers (among others) the stated requirements for these training situations. We have shown, that the three cases to a large extent rely on the requirements derived, while they still vary significantly in the way the game-designs make use of different role-based features and individualizations ARLearn offers. Consequently, the core contributions of this paper are the requirements gathered for multi-stakeholder decision training situations and their application in the developed cases.

Furthermore, we showed that ARLearn meets these requirements and thus appears to be a feasible environment for the design of according games. However, in the HeartRun case we where not able to provide a consistent, immersive gaming experience to our users. This problem motivated us to look for alternative interaction formats, which should not require the players to switch between device interaction and real-world activity.

Technological developments lead to the availability of wearable augmented reality glasses such as Google Glass[1]. Essentially, these devices comprise a wearable head-up-display, which projects images to the user, sensors detecting user location and orientation, network connection, camera, microphone and a touch panel. The device can be voice controlled, which allows for hands-free operation. These devices can also operate partly autonomously (i.e. react to a user's real world activity), once a specific app is started.

Consequently, we started to develop a new ARLearn client for Google Glass, complementing the existing mobile and web-based clients. The core idea of this client is to present events and messages caused by the game in the timeline of the player, allowing her to read messages, watch instruction movies or listen to event audios as part of the game play (compare Fig. 4). The information displayed on the screen is semi-transparent and the real world is still visible to the user.

[1] `http://www.google.com/glass/start/`

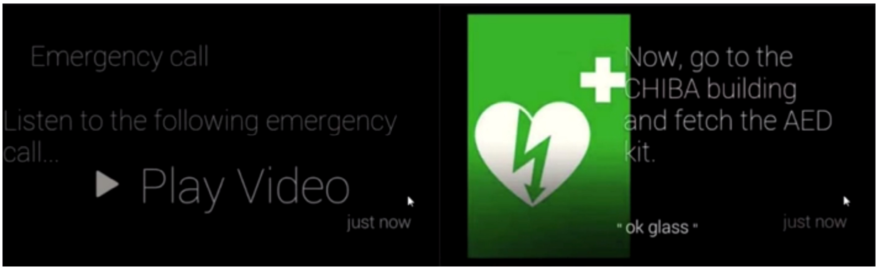



**Fig. 4.** Screenshots of the HeartRun game prototype for Google Glass

This way, the game play is not interrupted and players do not have to switch between device interaction and real world activity. Instructional videos, audio messages, or any other information can be received and watched even during performance of any real world activity (such as performing CPR to a training mannequin while watching the video how to it).

Our aim is to continue the developments of this ARLearn client in order to allow for full game interaction based on hands-free operations and pro-active information display via Google Glass. We believe, that we can provide a fully immersive game experience, which combines real world activities with game-play interactions in an augmented reality environment.

**Acknowledgements.** This research was co-financed by the European Regional Development Fund (ERDF), regions of the Euregio Meuse-Rhine under the INTERREG IVa program (EMR.INT4-1.2.-2011-04/070). The Dutch Institute of Advanced Logistics (DINALOG) sponsors the SALOMO project. The GLC at UNHCR funded the hostage taking game and gave valuable input to this work.

## References

1. Johnson, L., Levine, A., Smith, R., Stone, S.: The 2011 Horizon Report. The New Media Consortium, Austin (2011)
2. Carstens, A., Beck, J.: Get Ready for the Gamer Generation. Tech. Trends 49(3), 22–25 (2010)
3. Douch, R., Savill-Smith, C.: The Mobile Learning Network: The Impact of Mobile Game-Based Learning. In: Proc. of the IADIS Int'l Conf. Mobile Learning 2010, Porto, pp. 189–197 (2010)
4. Unterfauner, E., Marschalek, I., Fabian, C.M.: Mobile Learning With Marginalized Young People. In: Proc. of IADIS Int'l Conf. Mobile Learning 2010, Porto, pp. 28–36 (2010)
5. Liao, C.C.Y., Chen, Z.-H., Cheng, H.N.H., Chen, F.-C., Chan, T.-W.: My-Mini-Pet: A handheld pet-nurturing game to engage students in arithmetic practices. Journal of Computer Assisted Learning 27(1), 76–89 (2011), doi:10.1016/j.compedu.2011.01.009.
6. Douch, R., Attewell, J., Dawson, D.: Games technologies for learnig. More than just toys (2010), `http://goo.gl/mazqJ` (April 8, 2013)

7. Tichon, J., Hall, R.H., Hilgers, M.G., Leu, M., Agarwal, S.: Education and training in virtual environments for disaster management. In: Proc. of the World Conf. on Educational Multimedia, Hypermedia, and Telecommunications, Honolulu, Hawaii, pp. 1191–1194 (2003)
8. Haferkamp, N., Krämer, N.C.: Disaster Readiness through Education - Training Soft Skills to Crisis Units by means of Serious Games in Virtual Environments. In: Wolpers, M., Kirschner, P.A., Scheffel, M., Lindstaedt, S., Dimitrova, V. (eds.) EC-TEL 2010. LNCS, vol. 6383, pp. 506–511. Springer, Heidelberg (2010)
9. Dowell, J., Hoc, J.M.: Coordination in emergency operations and the tabletop training exercise. Le Travail Humain, 85-102 (1995)
10. Crichton, M.T., Flin, R., Rattray, W.A.: Training decision makers–tactical decision games. Journal of Contingencies and Crisis Management 8(4), 208–217 (2000)
11. Velasquez, J.D., Yoon, S.W., Nof, S.Y.: Computer-based collaborative training for transportation security and emergency response. Computers in Industry 61(4), 380–389 (2010)
12. Loiseau, M., Lavoué, E., Marty, J.-C., George, S.: Raising awareness on Archaeology: A Multiplayer Game-Based Approach with Mixed Reality. In: 7th European Conference on Games Based Learning (ECGBL 2013), Porto, Portugal, October 3-4, pp. 336–343 (2013)
13. Soar, J., Mancini, M.E., Bhanji, F., Billi, J.E., Dennett, J., Finn, J., Morley, P.T.: Part 12: Education, implementation, and teams: 2010 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science with Treatment Recommendations. Resuscitation 81, e288(2010)
14. 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.) Proc. of 11th World Conference on Mobile and Contextual Learning (mLearn 2012), Helsinki, Finland, October 16-18, pp. 244–247 (2012)
15. Kurapati, S., Kolfschoten, G., Verbraeck, A., Drachsler, H., Specht, M., Brazier, F.: A Theoretical Framework for Shared Situational Awareness in Sociotechnical Systems. In: Proc. 2nd Workshop on Awareness and Reflection in Technology-Enhanced Learning, pp. 47–53 (2012)
16. de Bruijn, H., ten Heuvelhof, E.: Management in networks: On multi-actor decision making, 1st edn. Routledge, Oxford (2008)
17. Salmon, P.M., Stanton, N.A., Walker, G.H., Jenkins, D.P., Mcmaster, R., Young, M.S.: What really is going on? Review of situation awareness models for individuals and teams. Theoretical Issues in Ergonomics Science, 37–41 (2008)
18. 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)
19. Lave, J., Wenger, E.: Situated learning: Legitimate peripheral participation. Cambridge University Press (1991)
20. De Freitas, S.: Learning in immersive worlds, A review of game-based learning. Joint Information Systems Committee, London (2006), `http://goo.gl/ob2B1I`
21. Marchiori, E.J., Ferrer, G., Fernandez-Manjon, B., Povar-Marco, J., Suberviola, J.F., Gimenez-Valverde, A.: Video-game instruction in basic life support maneuvers. Emergencias 24, 433–437 (2012)
22. Crichton, M., Flin, R.: Training for emergency management: tactical decision games. Journal of Hazardous Materials 88(2), 255–266 (2001)
23. Crichton, M.T.: Improving team effectiveness using tactical decision games. Safety Science 47(3), 925–7535 (2009) ISSN 0925-7535

24. Bikovska, J., Merkuryeva, G.: Scenario-Based Planning and Management of Simulation Game: A Review. In: 21st European Conf. on Modelling and Simulation (2007)
25. Ponder, M., Herbelin, B., Molet, T., Schertenlieb, S., Ulicny, B., Papagiannakis, G., Magnenat-Thalmann, N., Thalmann, D.: Immersive VR decision training: Telling interactive stories featuring advanced virtual human simulation technologies. In: Proc. of the Workshop on Virtual Environments 2003, pp. 97–106. ACM (2003), `http://goo.gl/6qcEMJ`
26. Chittaro, L., Ranon, R.: Serious games for training occupants of a building in personal fire safety skills. In: Proc. of VS-GAMES 2009: IEEE First Int'l Conf. on Games and Virtual Worlds for Serious Applications, Coventry, UK, pp. 76–83 (2009)
27. Dobson, M., Kyrylov, V., Kyrylova, T.: Decision Training using Agent-based Business Strategy Games. In: CATE, pp. 66–71 (2004)
28. Ternier, S., Klemke, R., Kalz, M., Van Ulzen, P., Specht, M.: ARLearn: Augmented reality meets augmented virtuality. J. of Universal Computer Science: Special Issue on Technology for Learning Across Physical and Virtual Spaces 18(15), 2143–2164 (2012)
29. Ternier, S., De Vries, F., Börner, D., Specht, M.: Mobile augmented reality with audio. Supporting fieldwork of Cultural Sciences students in Florence. In: InSuEdu 2012 Workshop at the SEFM 2012, the 10th International Conference on Software Engineering and Formal Methods, Thessaloniki, Greece (2012)
30. Van Rosmalen, P., Klemke, R., Westera, W.: Alleviating the Entrance to Serious Games by Exploring the Use of Commonly Available Tools. In: Proceedings of the 5th European Conference on Games Based Learning, Athens, October 20-21, pp. 613–619 (2011)
31. Schmitz, B., Ternier, S., Kalz, M., Klemke, R., Specht, M.: Designing a mobile learning game to investigate the impact of role-playing on helping behaviour. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 357–370. Springer, Heidelberg (2013)
32. Miao, Y., Hoeksema, K., Hoppe, H.U., Harrer, A.: CSCL scripts: Modelling features and potential use. In: Proc. of the 2005 Conf. on Computer Support for Collaborative Learning, Int'l Society of the Learning Sciences, pp. 423–432 (2005)
33. Dillenbourg, P., Hong, F.: The mechanics of CSCL macro scripts. International Journal of Computer-Supported Collaborative Learning 3(1), 5–23 (2008)
34. Kelle, S., Klemke, R., Specht, M.: Design patterns for learning games. International Journal of Technology Enhanced Learning 3(6), 555–569 (2011)
35. Gagnon, D.J.: ARIS: An open source platform for developing mobile learning experiences. University of Wisconsin at Madison, Madison (2010), `http://goo.gl/ycI7fi`
36. Holden, C.L., Sykes, J.M.: Leveraging mobile games for place-based language learning. International Journal of Game-Based Learning (IJGBL) 1(2), 1–18 (2011)
37. 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)
38. Robles, G., Gonzales-Barahona, J.M., Fernandez-Gonzales, J.: Implementing Gymkhanas with Android smartphones: A multimedia m-learning game. In: 2011 IEEE Global Engineering Education Conf (EDUCON), pp. 960–968. IEEE (April 2011)

# The Group Formation Problem: An Algorithmic Approach to Learning Group Formation

Johannes Konert, Dmitrij Burlak, and Ralf Steinmetz

Technische Universität Darmstadt, Germany
{johannes.konert,ralf.steinmetz}@kom.tu-darmstadt.de,
dmitrij.burlak@gmail.com

**Abstract.** Fostering knowledge exchange among peers is important for learners' motivation, achievement of learning goals as well as improvement of problem solving competency. Still, the positive effects of such an exchange depend strongly on the suitability of the selected peers in a group.

A comparison of existing algorithmic solutions applicable for E-learning and CSCL scenarios reveals limited support for requirements derived from related work in pedagogical psychology. Therefore, the GroupAL algorithm is proposed. It supports the use of criteria that are either expected to be matched homogeneous or heterogeneous among participants while aiming for equally good group formation for all groups. A normed metric allows for comparison of different group formations and is robust against variations. Finally, the evaluation reveals the advantages and widespread applicability of GroupAL. Compared to existing solutions, it achieves a better group formation quality under the chosen conditions.

**Keywords:** Learning Group Formation, Algorithmic Optimization, Matching, CSCL.

## 1 Introduction and Motivation

The didactic concept of collaboration in small working groups is especially well suitable for tasks aiming for agency of problem solving competency [2]. For development of this increasingly important competency and a successful handling of open-format problems peers discuss their point of view and favored approaches to a problem. Open-format problems are characterized by missing pre-defined approaches or only single correct solutions. While working on such problems, peers can complement one another in learning style and benefit from advances in knowledge or skills. If this potential is used, peers act as an integrated group on analysis and solution of the problem at hand [3].

Fruitful cooperation in the group depends on group size, suitable open-format problems, and manifold other criteria which are related to capabilities and characteristics of the learners, as well as criteria concerning learning context and group constellation. If there is a misalignment within the group concerning

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 221–234, 2014.
© Springer International Publishing Switzerland 2014

these criteria, solo attempts as well as digression and missing motivation of individual group members occur [21]. Generally, in classroom scenarios (physical presence of learners) the pedagogue or seminar lecturer knows the learners and can build learning groups manually while considering the criteria. Studies revealed a strong, disadvantageous influence of friendship among the learners and contiguous seats in the moment of group allocation [22]. Thereby, the diverse preconditions of learners, as mentioned above, are not sufficiently considered in the group formation process that aims for maximizing knowledge gain of each individual and optimizing mutual assistance in problem analysis and solution development.

Motivation to develop a suitable, algorithm-based method providing assistance to lecturers for sophisticated learning group formation is further increased, if learners are spread over diverse locations and act as an inter-connected community. If such E-learning environments focus on support for self-regulated learning and do not enforce specific amounts of time invested or do not offer preset learning goals, diversity in preconditions and learning targets of participating learners are (more) diverse [11]. In case of unsupervised learning environments or in case of a higher number of participants that exceeds a certain (classroom typical) limit, manual group formation by lecturers is impractical and algorithmic solutions are not only of great support but strongly needed.

> *The challenge to optimize learning group formation from a given set of peers to match, while respecting homogeneously to match criteria simultaneously with heterogeneously to match criteria and aiming for a balanced quality of the build groups, is called the* Group Formation Problem*.*
> [17, p. 16]

None of the currently existing algorithmic solutions aiming for assistance in such learning group formation incorporates the mentioned aspects sufficiently (see section 2 below). As illustrated above, such a solution is expected to allow manifold criteria, with individual weighting and dimensionality. Additionally, flexibility is desired that allows combination of criteria which are matched homogeneously with criteria which are matched heterogeneously among learners of one group.

In the following, metrics are derived for evaluation of learning group formation quality allowing for the desired combination of criteria. The metrics are used as a basis for the proposed group formation algorithm GroupAL. The results of the conducted evaluation study prove the benefits and improved group formation quality compared to the existing solutions from related work under the chosen conditions.

The contribution at hand is a revised, condensed and translated overview based on earlier publications concerning GroupAL [18, 19, 17].

## 2 Related Work

Collaboration in small (learning) groups, work phases, and their beneficial conditions, are investigated in the field of pedagogical psychology since the Sixties

[30]. Since then, manifold research studies confirmed the positive effects of collaboration, group-based learning and knowledge exchange among peers. Therefore, Damon [7] concluded for the field of didactic that exchange among peers is a suitable amendment for any teaching scenario. One major prerequisite is the suitability of the problem definition to be addressed by multiple peers collaboratively. It this is not fulfilled, collaboration can still be achieved by knowledge exchange and mutual feedback which can be improved considerably by a structured moderation (of feedback) and preset interaction patterns [29].

After provision (by instructors) or selection (by learners) of problem definition, learning place(s) and interaction patterns, learning groups can be formed under consideration of the context. In case of differences in level of knowledge and abilities among learners, it appears obvious to elect some of them as tutors who pass on their knowledge advances to others in the group (or to a peer in learning tandems) [15, 31]. Even though *learning by teaching* appears to be advantageous likewise for the tutor, contrary argumentation exist. Based on Piaget's constructivist didactic, Damon [7, p. 334] argues mutual respect to be a basic requirement for social interaction and knowledge exchange among learners. But this respect is compromised by a too significant difference in the level of knowledge and abilities. This insight motivates the intention to match peers in learning groups that support mutual amendment based on diversity of knowledge while the level of knowledge is equal (*symmetry of knowledge*) [9, p. 7]. A suitable algorithmic representation model to map (part of) knowledge areas to learning problems and learning targets is given by the skill tree structure of the *knowledge spaces* proposed by Albert and Lukas [1]. Usage of such structures aims for coverage of all related skill tree areas by all learners after solving the provided learning problems. Beside criteria related to the knowledge (structures), learning style preferences are expected to be matched heterogeneously to allow group members to share their (different) views. The resulting cognitive dissonances lead to argumentation and exchange of diverging approaches and considerations, finally resulting in a more comprehensive work on the field of the provided learning problem and inter-relations [8]. Constitutional theories and models for evaluation of learning styles and subsequent group formation are proposed and evaluated by Felder and Silverman [10] or Kolb and Kolb [16]. Additional criteria related to the individual learner are age, gender, geo-location and intensity of work (working hours), which are generally matched homogeneously within a learning group.

Besides criteria related to the individual leaner, group-related criteria are also relevant. Most dominant aspect is the finding of the optimal size of a learning group. Even though the group size depends on learning problem characteristics and expected duration of collaboration within the group, studies identified the optimal group size between three to six learners [27].Such a size allows for adequate exchange and eventually desired heterogeneity of specific criteria within the group while avoiding emergence of sub-groups and redundancy which may lead to exclusion of group members in the end. In addition, attention to group roles and responsibility for parts of the addressed problem, is conducive for exchange within the learning group [20].

Extensive models for learning group formation emerged in the field of pedagogical psychology [6, 21, 26]. The related publications name further aspects and details concerning interplay of group members in learning groups. These aspects can be considered and weighted by users of GroupAL. Still, for the further algorithmic examination of the group formation problem these additional aspects are not discussed in detail here.

## 3 Goals for Algorithmic Learning Group Formation

In summary, the analysis of related work suggests the following four goals to be achieved by the desired method for algorithmic learning group formation:

**G1** extendable modeling, exchangeability, and weighting of criteria used for group formation; due to the fact that no generally accepted list of criteria exists and criteria to be used depend on the context the groups are build for;
**G2** support for the creation of homogeneous, heterogeneous and mixed learning groups in several criteria simultaneously; due to the beneficial influence of the *symmetry of knowledge* and mutual amendment;
**G3** assessment and optimization of group formations based on a group formation metric that takes into consideration the constellation of group members; due to the targeted mutual exchange and avoidance of exclusion of members;
**G4** minimization of the differences among the formed learning groups; due to acceptance by learners and fairness for all participants, e. g. in classroom scenarios.

## 4 State of the Art in Algorithmic Learning Group Formation

The identified existing algorithmic solutions for learning group formation can generally be separated into two groups of approaches: *semantic matchmakers* and *non-linear optimization techniques.*

*Semantic matchmakers* utilize ontologies to calculate how well two (or more) learners match for optimal achievement of the set learning goals. Ontologies allow precise formulation for extensive boundary conditions to be respected during group formation [13]. However, application of ontology-based approaches becomes costly, if no suitable ontology exists or boundary conditions – as desired by lecturers – are not directly expressible. Moreover, semantic matchmakers unfortunately do not express the quality of created learning groups in a comparable value and do not include the desired equable distribution of learning group formation quality among all created groups. Still, semantic matchmakers are

**Table 1.** Approaches to algorithmic learning group formation; • agent system, evaluates suitability of candidates iteratively based on task solutions for a selected homogeneous or heterogeneous strategy, ° statement about restriction violations, * using a threshold, ⋆ using a heuristic.

| *System* | Calculation of Group Formation Quality | Uniform Group Formation Quality | Infinite Number of Criteria | Criteria Weighting | Several Algorithms Available | Homogeneous Group Formation | Heterogeneous Group Formation | Mixed Group Formation |
|---|---|---|---|---|---|---|---|---|
| *Fits/CL* [13] | - | + | - | - | - | + | + | + |
| *GroupMe* [23] | +° | - | + | + | - | + | + | + |
| *I-minds* [28] | -• | - | - | - | - | + | + | -• |
| *GroupFormation* [5] | - | - | - | - | - | + | + | - |
| *Together* [25] | +* | - | - | - | - | - | + | - |
| *OmadoGenesis* [12] | + | - | + | - | + | + | + | + |
| *TeamMaker* [4] | +⋆ | - | + | + | - | + | + | + |

well suited for large E-learning scenarios and for expressing complex dependencies to be considered for matchmaking. The benefits and drawbacks of semantic matchmakers and ontoloty-based group formation are derived from two semantic matchmakers whose characteristics are listed for comparison in the first rows of Table 1.

*Non-linear optimization techniques* use a representation of the desired person-criteria as an $n$-dimensional feature space (vector) for each learner. Group-related criteria are respected as boundary conditions or within the metric that is used to calculate the group formation quality (*objective quality function*). Based on the feature spaces as input, cluster analysis can be used to match similar learners in respect to the homogeneously to match criteria (similarity). Such an approach can be implemented, e. g. using Fuzzy-C-Means [25]. Unfortunately, this approach is limited in case both, heterogeneously and homogeneously to match criteria, need to be respected during group formation. Here heuristics and iterative optimization can be used as demonstrated by Cavanaugh and Ellis [4]. Only a limited number of solutions is based on algorithms which are specifically developed to address the stated requirements and go beyond classic optimization techniques. Generally, non-linear optimization techniques are used in smaller E-learning scenarios and web-based systems with a limited number of considerable criteria. The analyzed systems, based on non-linear optimization, are listed in the lower part of Table 1 for comparison.

The tabular disposition of existing approaches in Table 1 reveals the importance to focus the GroupAL development on support for homogeneously and heterogeneously to match criteria simultaneously, allow criteria weighting and calculation of a normed quality metric that allows comparison of several build group cohorts (see goals G1-G4).

For comparison the following sections focus on the non-linear optimization techniques which allow group formation with respect to only heterogeneously to match criteria or homogeneously and heterogeneously to match criteria simultaneously (GroupFormationTool [5], OmadoGenesis [12], Together [25], and TeamMaker [4]). Techniques focusing on support for homogeneous criteria only are primarily based on established clustering approaches and are covered in other publications [14]. Semantic matchmakers are not focused in the following as they are based on ontologies and thus applicability is less flexible in case criteria need to be easily exchangeable by end-users (e. g. instructors).

# 5 GroupAL Group Formation Algorithm

To achieve the four goals (G1-G4) as listed above, a metric will be defined in this section measuring the quality of a whole cohort of created learning groups in the interval $(0, 1)$. First, basic definitions will be introduced, then a metric to calculate the suitability of two participants is presented (*PairPerformanceIndex PPI*). It builds the basis to define a metric for one group (*GroupPerformanceIndex GPI*) and as a final step a metric for the cohort of all groups (*CohortPerformanceIndex CPI*). The CPI is calculated in order to ensure the minimization of differences between groups as described for goal G4.

Definitions are derived from Ounnas et al. [24] and are extended with a focus on the elements necessary for the definition of the group formation quality metric.

## 5.1 Basic Definitions

**Criteria.** A criterion is defined as a vector $k \in \mathbb{R}^n$, which is considered to be used as a relevant parameter, variable or characteristic for group formation. The set $K$ of criteria is finite. $K = \{\{k_1, k_2, \ldots, k_q\} | \forall j = 1, \ldots, q,\ k_j \in \mathbb{R}^n\}$.

**Disjunctive Criteria Sets.** A criterion has to be assigned explicitly to one of the following two disjunctive sets. A criterion is *homogeneous*, if the criterion's value should be preferably similar in a build group ($K_{hom}$). On the contrary, a *heterogeneous* criterion is expected to result in amendatory values among the group members ($K_{het}$). Because they are disjunctive, $K_{hom} \cap K_{het} = \varnothing \wedge K_{hom} \cup K_{het} = K$.

**Participants.** The finite set of participants is defined as $P = \{p_1, p_2, \ldots, p_M\}$. Each participant is characterized by a set of criteria $p \subseteq K$. The criteria used for comparison need to be equal for all participants. $M = |P| > 1$ is the number of participants.

**Groups.** A finite set of participants $p \in P$ is defined as a group $g$, if it has at least 2 elements $|g| > 1$ (minimal group). One element $p_i \in g$ is called a *member* of the group. $G_x \subseteq G$ is defined as the set of all groups with a fixed size $X$. Consequently, group cardinality is $\mathrm{N} = \frac{M}{X} \forall X \geq 1$.

**Cohorts.** A cohort $C$ is a set of pairwise disjoint groups $g_1, g_2, \ldots, g_s$: $\forall p \in P \neg \exists g_1, g_2 \in G : p \in g_1 \wedge p \in g_2$. A cohort contains all participants. Additionally, a cohort only consists of groups with the same fixed size $X$.

### 5.2 Defining Pair Performance Index (PPI)

The PPI uses a weighted normalized distance function ($wd$) as a basis where each criterion can have an individual weight. The*normalized Manhattan distance* is used as underlying distance function ($d$) to calculate how similar two participants in the values of one criterion are (see Equation 1).

$$\begin{aligned} &wd : [0,1]^n \times [0,1]^n \times [0,1] \rightarrow [0,w], \\ &wd\left(k_p^1, k_p^2, w_p\right) = w_p * d\left(k_p^1, k_p^2\right), \end{aligned} \tag{1}$$

where $k_p^1$ and $k_p^1$ are criterion vectors for one criterion of two participants, $n$ the dimensionality of $k_p$, $w_p$ the weight for this criterion with $w_p \in [0,1]$ and the sum of all weights $\sum_{t=1}^{q} w_t = 1$. In contrast to e. g. the *Euclidean* distance, it is a linear function, appropriate to express how complete the dimension space of a criterion is covered by two participants. This is of particular importance for heterogeneously to match criteria (see Equation 2, *homSum* is calculated analogue for all $k_i \in K_{hom}$).

$$\begin{aligned} &hetSum : K \times K \times \{0,1\}^n \rightarrow \left[0, \sum\nolimits_{i=1}^{|K_{het}|} w_i\right] \in [0,1], \\ &hetSum\left(K_{het}^1, K_{het}^2, W\right) = \sum\nolimits_{i=1}^{|K_{het}|} wd\left(k_i^1, k_i^2, w_i\right), \end{aligned} \tag{2}$$

where $|K_{het}|$ is the set of heterogeneous criteria, and $K_{het}^1$ and $K_{het}^2$ are value vectors of the criteria for two participants.

Hence, the PPI is calculated as the sum of distances for all heterogeneously to match criteria (*hetSum*) minus the sum of all distances of homogeneously to match criteria (*homSum*) as shown in Equation 3. Consequently, the PPI reaches its maximum in case distances for homogenous criteria is ideally zero and for heterogeneous criteria the whole $n$-dimensional space of each criterion is covered. As such, and with the possibility to weight criteria, the PPI fulfills goal G1 stated above and delivers a solutions for G2 (up to here, limited to group size of two participants).

$$\begin{aligned} &PPI : K \times K \times \{0,1\}^n \rightarrow \left[-\sum\nolimits_{i=1}^{|K_{hom}|} w_i, \sum\nolimits_{j=1}^{|K_{het}|} w_j\right] \in [-1,1], \\ &PPI\left(K^1, K^2, W\right) = hetSum\left(K_{het}^1, K_{het}^2, W\right) - homSum\left(K_{hom}^1, K_{hom}^2, W\right) \end{aligned} \tag{3}$$

For better usability in the following, PPI will be normalized as NPPI $\in [0,1]$ by shifting the *PPI*-value by the maximum negative value (the maximum value of *homSum*) and division by the resulting maximum value (sum of maximal value of *homSum* plus maximum value of *hetSum*) which equals to $\sum_{t=1}^{|W|} w_i$ which equals to 1. Ideally, no division is necessary. The division is kept in Equation 4 to allow the GroupAL algorithm to cope with liberalization (sum of weights $\neq 1$).

$$\begin{aligned} &NPPI : K \times K \times \{0,1\}^n \rightarrow [0,1], \\ &NPPI\left(K^1, K^2, W\right) = \frac{PPI\left(K^1, K^2, W\right) + \sum_{i=1}^{|K_{hom}|} w_i}{\sum_{t=1}^{|W|} w_i} \end{aligned} \tag{4}$$

### 5.3 Defining Group Performance Index (GPI) and Cohort Performance Index (CPI)

Concerning goal G2 for arbitrary group size and to calculate a metric how well participants in a group match altogether, the mean value of all possible $\binom{X}{2}$ NPPIs in a group is calculated ($\overline{NPPI}$). A mean value is not sufficient to respect disadvantageous constellations, e. g. deviators. Consequently, GPI will as well take the normalized standard deviation of the group's NPPIs into account as shown in Equation 5. The fewer isolated individual participant in the group exist, the higher the overall GPI value is as requested with goal G3. The same approach is used to calculate the quality of a complete cohort of groups (CPI). It is the product of mean GPI ($\overline{GPI}$) multiplied with the normalized standard deviation of all GPIs in the cohort (see Equation 6). Consequently, if groups have dissimilar GPI values, it results in a low CPI as requested with goal G4. Beside normalized standard deviation other variation methods could have been taken into account as discussed by Konert [17, p. 79].

$$\begin{aligned} &GPI : G \rightarrow [0,1], \\ &GPI(g) = \overline{NPPI} * \left(\frac{1}{1 + \sigma_{NPPIs}}\right) \end{aligned} \tag{5}$$

$$\begin{aligned} &CPI : C \rightarrow [0,1], \\ &CPI(c) = \overline{GPI} * \left(\frac{1}{1 + \sigma_{GPIs}}\right) \end{aligned} \tag{6}$$

### 5.4 The GroupAL Matcher Algorithms

The matching algorithms use the defined metrics (PPI, GPI and CPI) to assign participants one by one to learning groups until all participants are assigned. Initially, $N$ empty groups are created and all participants in set $P$ are added to the group of not matched participants (*NMP*). Each group is assigned a random pivot element from *NGP*. Essentially, two different matching strategies where implemented for GroupAL as the matching approach can mainly influence the quality of achieved results.

**Group-Centric Matcher.** The *Group-Centric Matcher (GCM)* selects a random group first and then moves the one candidate from *NGP* into the group that increases the resulting *GPI* of the group the most on a percentage basis. This addition of the best candidate is continued until the group reaches it's targeted size of $X$ members. The algorithm continues with the next group until all groups are processed or $NMP = \emptyset$. *GCM*'s behavior is expressed formally in Equation 7.

$$\left\{ g_{fix} \cup p \middle| \; \forall p \in NMP, g_{fix} \in G_x : |g_{fix}| < X \wedge max \left( \frac{GPI\left(g_{fix} \cup t\right)}{GPI\left(g_{fix}\right)} \right) \right\} \quad (7)$$

**Participant-Centric Matcher.** In variation to *GCM*, the *Participant-Centric Matcher (PCM)* selects randomly candidate by candidate from the set *NMP* and moves them into the group whose *GPI* is increased the most by addition of this candidate on a percentage basis. The *PCM*'s behavior is expressed in Equation 8 for comparison.

$$\left\{ g \cup t_{fix} \middle| \; p_{fix} \in NMP, \forall g \in G_x : |g| < X \wedge max \left( \frac{GPI\left(g \cup p_{fix}\right)}{GPI(g)} \right) \right\} \quad (8)$$

## 6 Evaluation

To measure effectiveness in group formation, GroupAL Matchers (*GCM* and *PCM*) are compared to the solutions GroupFormation [5], Together [25], OmadoGenesis [12], and TeamMaker [4].

### 6.1 Study Design

The comparison of the matchers uses two different data setups, each providing the matching algorithms with generated data sets containing 500 participants. Setup $\alpha$ (S$\alpha$) contains for each participant 1 criterion $k_{het,1} \in K_{het} \mid dim(k_{het,1}) = 4$ and is used for comparison with the algorithms of related work that are only capable to process one criterion or support only a maximum of 4 dimensions (GroupFormation, Together, OmadoGenesis). Setup $\beta$ (S$\beta$) contains for each participant 4 criteria $k_{het,1}, k_{het,2} \in K_{het} \wedge k_{hom,1}, k_{hom,2} \in K_{hom} \mid dim(k_i) = 4 \; \forall i \in [0, 1, \ldots, 4] \wedge k_i \in K$. This setup is used for comparison with TeamMaker whose capabilities come most close to GroupAL as it supports several homogeneous and heterogeneous criteria simultaneously, too.

Orthogonally, for setup $\alpha$ three variations of participants' criteria value distribution have been generated to investigate robustness regarding differences in value distribution (V1 even distribution, V2 normal distribution, V3 evenly distributed extreme values (only the values 0 and 1). As TeamMaker is designed for discrete criteria values, in setup $\beta$ only even distribution of extreme values has been generated (V3).

To eliminate random effects, the four data sets (S$\alpha$ V1-V3, S$\beta$ V3) were generated 100 times and used for 100 runs of the matching algorithms. In each run the matchers were started three times to create group formatios with group sizes of 2,3, and 6 members.

### 6.2 Results and Interpretation

For setup $\alpha$ the average *CPI* values of all 100 runs for each value distribution variation (V1-V3), itemized for the three different targeted learning group sizes (2,3,6), is visualized in Figure 1.

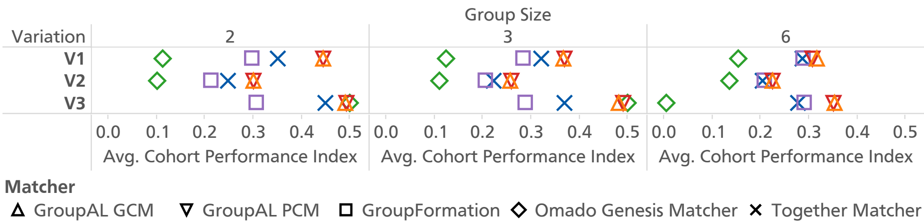



**Fig. 1.** Matcher differences in setup $\alpha$

For setup $\beta$ the average *CPI* and average *GPI* values of all 100 runs for the evenly distributed extreme values (V3), itemized for the three different targeted learning group sizes (2,3,6), is visualized in Figure 2. The plot shows two different calculation for *CPI* and *GPI*: on the left side based on the GroupAL quality metric as derived above and on the right side as defined and calculated by TeamMaker. This comparison allows to prove that GroupALs group formations are competitive even if the quality metric of TeamMaker is used. *GPI* and *CPI* calculations are not normalized by TeamMaker, resulting in negative values as the scale depends on the criteria and weights [4, p.8].

*Interpretation.* As clearly visible, in both setups the GroupAL matchers (*GCM* and *PCM*) achieve higher *CPI* values under the chosen conditions compared to the algorithms from related work which were compared here. Concerning the above stated goals G1-G4 and the derived quality metric for *GPI* and *CPI*, it is reasonable to conclude that GroupAL has a higher capability of matching participants with respect to diverse criteria combination while aiming for equated formation quality both, within groups (*GPI*), and of all groups within the resulting cohort (*CPI*).

In setup $\alpha$ on extreme criteria values (V3), OmadoGenesis achieves slightly higher *CPIs* for small group sizes (2 and 3 members), but fails to match adequate candidates for groups with 6 members. In all other variations of setup $\alpha$

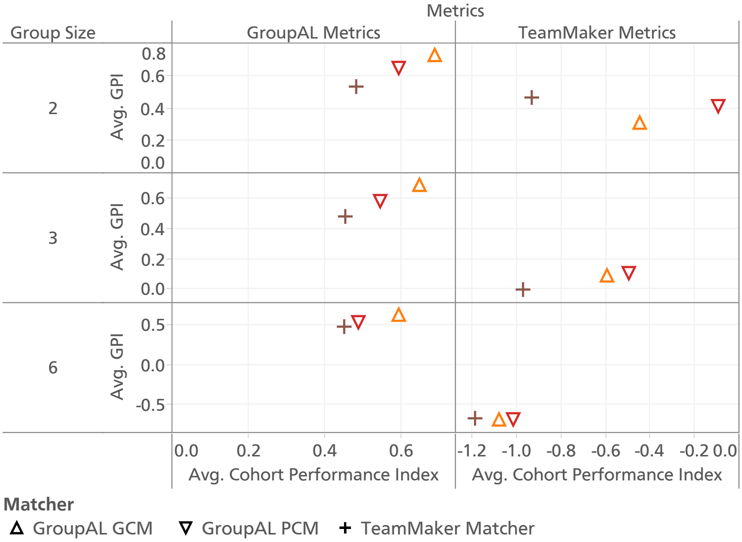



**Fig. 2.** Matcher differences in setup $\beta$

GroupAL's *GCM* and *PCM* deliver better results which are very close to each other. Only for groups with 6 members on evenly distributed criteria values (V1), GroupAL's *GCM* achieves slightly better results than GroupAL's *PCM*. Both GroupAL matchers appear to be quite robust against variations in criteria value distribution and requested group sizes.

In setup $\beta$ the benefits of a normalized *GPI* and *CPI* can clearly be seen on the right side (GroupAL metrics). On the left side the results reveal that GroupAL matchers still achieve higher quality metric values even though *GPI* and *CPI* were calculated by the quality functions of TeamMaker.

## 7 Conclusion and Outlook

The discussion and revision of relevant criteria for group formation, as identified by related work from pedagogical psychology, lead to the conclusion that four goals need to be achieved by algorithmic learning group formation (G1-G4) in order to address the *group formation problem*. Beside other aspects, the demand for easy exchangeability of heterogeneously and homogeneously to match criteria or weights by instructors, the necessity to aim for evenly distributed group formation quality, and the objective to achieve beneficial exchange for all participants in a group, lead to the conclusion that non-linear optimization is preferable compared to semantic, ontology-based approaches.

After examination of existing algorithmic solutions and discussion of identified limitations, the GPI and CPI metrics of GroupAL were derived and defined formally, respecting the demanded goals G1-G4. The simulative evaluation of the two GroupAL matchers (*GCM* and *PCM*) revealed their capability to form learning groups that achieve significantly higher Group Performance (*GPI*) and Cohort Performance (*CPI*) results as the compared algorithms from related work under the chosen conditions.

The implementation of GroupAL matchers (*GCM* and *PCM*) and evaluation metrics (*PPI*, *GPI*, and *CPI*) will soon be released under an Open-Source license. As data sets for comparison of learning group formation algorithms are very limited, the generated data sets will be included. Interested parties are advised to search under the term 'GroupAL' regularly or visit the authors' websites for further information.

The advantages of GroupAL could further be corroborated by a field study to prove that learners indeed benefit significantly from peers in groups build based on GroupAL, i.e. *GPI* metric. Thus, future research includes the integration of GroupAL into E-learning environments like e.g. Moodle[1] to conduct long-term studies comparing on one side the effectivity based on achieved learning outcomes of participants in groups build by GroupAL and on the other side groups that are build randomly or by the participants themselves.

## References

1. Albert, D., Lukas, J.: Knowledge Spaces: Theories, Empirical Research, and Applications. Psychology Press (1999)
2. Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.J., Weiß, M.: Erfassung fächerübergreifender Problemlösekompetenzen in PISA. Tech. rep. OECD PISA Deutschland, Berlin (1999)
3. Borsch, F.: Kooperatives Lehren und Lernen im schulischen Unterricht, 1st edn. Kohlhammer, Stuttgart (2010)
4. Cavanaugh, R., Ellis, M.: Automating the Process of Assigning Students to Cooperative-Learning Teams. In: Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition (2004)
5. Christodoulopoulos, C.E., Papanikolaou, K.A.: A Group Formation Tool in an E-Learning Context. In: 19th IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007), October 2007, pp. 117–123 (2007)
6. Cohen, E.G., Goodlad, J.I.: Designing Groupwork: Strategies for the Heterogeneous Classroom, 2nd edn. Teachers College Press (1994)
7. Damon, W.: Peer Education: The Untapped Potential. Journal of Applied Developmental Psychology 5(4), 331–343 (1984)
8. de Los Angeles Constantino-González, M., Suthers, D.D., Escamilla De Los Santos, J.D.: Coaching Web-based Collaborative Learning based on Problem Solution Differences and Participation. International Journal of Artificial Intelligence in Education 13(2-4), 263–299 (2003)
9. Dillenbourg, P.: What do you mean by Collaborative Learning? In: Dillenbourg, P. (ed.) Collaborative-learning: Cognitive and Computational Approaches, pp. 1–15. Elsevier, Oxford (1999)
10. Felder, R.M., Silverman, L.K.: Learning and Teaching Styles. Engineering Education 78, 674–681 (1988)

[1] `http://www.moodle.org`, last visited on 04/01/2014

11. Garton, L., Haythornthwaite, C., Wellman, B.: Studying Online Social Networks. Journal of Computer-Mediated Communication 3(1), 1–9 (2006)
12. Gogoulou, A., Gouli, E., Boas, G., Liakou, E., Grigoriadou, M.: Forming Homogeneous, Heterogeneous and Mixed Groups of Learners. In: Brusilovsky, P., Grigoriadou, M., Papanikolaou, K. (eds.) Proceedings of Workshop on Personalisation in E-Learning Environments at Individual and Group Level, 11th International Conference on User Modeling, pp. 33–40 (2007)
13. Inaba, A., Supnithi, T., Ikeda, M., Mizoguchi, R., Toyoda, J.: How Can We Form Effective Collaborative Learning Groups? In: Gauthier, G., VanLehn, K., Frasson, C. (eds.) ITS 2000. LNCS, vol. 1839, pp. 282–291. Springer, Heidelberg (2000)
14. Jain, A.K., Murty, M.N., Flynn, P.J.: Data Clustering: A Review. ACM Computing Surveys 31(3), 264–323 (1999)
15. Kester, L., Van Rosmalen, P., Sloep, P., Brouns, F., Kone, M., Koper, R.: Matchmaking in Learning Networks: Bringing Learners Together for Knowledge Sharing. Interactive Learning Environments 15(2), 117–126 (2007)
16. Kolb, A.Y., Kolb, D.A.: The Kolb Learning Style Inventory - Version 3.1 Technical Specifications. Tech. rep. HayGroup, Boston, USA (2005)
17. Konert, J.: Interactive Multimedia Learning: Using Social Media for Peer Education in Single-Player Educational Games (accepted for publication). Ph.d. thesis, Technische Universität Darmstadt, Germany (2014)
18. Konert, J., Burlak, D., Göbel, S., Steinmetz, R.: GroupAL: Ein Algorithmus zur Formation und Qualitätsbewertung von Lerngruppen in E-Learning-Szenarien mittels n-dimensionaler Gütekriterien. In: Breitner, A., Rensing, C. (eds.) Proceedings of the DeLFI 2013: Die 11. e-Learning Fachtagung Informatik der Gesellschaft für Informatik e.V., pp. 71–82. Köllen, Bremen (2013)
19. Konert, J., Burlak, D., Göbel, S., Steinmetz, R.: GroupAL: ein Algorithmus zur Formation und Qualitätsbewertung von Lerngruppen in E-Learning-Szenarien. i-com 13(1), 70–81 (2014)
20. Lisak, A., Erez, M.: Leaders and followers in multi-cultural teams. In: Proceeding of the 2009 International Workshop on Intercultural Collaboration, IWIC 2009, p. 81. ACM Press, New York (2009)
21. Michaelsen, L.K., Fink, L.D., Hall, A.: Designing Effective Group Activities: Lessons for Classroom Teaching and Faculty Development. In: DeZure, D. (ed.) To Improve the Academy: Resources for Faculty, Instructional and Organizational Development, New Forums, Stollwater (1997)
22. Mitchell, S.N., Reilly, R., Bramwell, F.G., Lilly, F.: Friendship and Choosing Groupmates: Preferences for Teacher-selected vs. Student- selected Groupings in High School Science Classes. Journal of Instructional Psychology 31(1), 1–6 (2012)
23. Ounnas, A., Davis, H., Millard, D.: A Framework for Semantic Group Formation. In: Eighth IEEE International Conference on Advanced Learning Technologies, pp. 34–38 (2008)
24. Ounnas, A., Millard, D.E., Davis, H.C.: A Metrics Framework for Evaluating Group Formation. In: Proceedings of the 2007 International ACM Conference on Supporting Group Work, GROUP 2007, p. 221 (2007)
25. Paredes, P., Ortigosa, A., Rodriguez, P.: A Method for Supporting Heterogeneous-Group Formation through Heuristics and Visualization. Journal of Universal Computer Science 16(19), 2882–2901 (2010)
26. Race, P.: Making Learning Happen, vol. 2. Sage Publications, London (2010)
27. Shim, K.J., Srivastava, J.: Team Performance Prediction in Massively Multiplayer Online Role-Playing Games (MMORPGs). In: 2010 IEEE Second International Conference on Social Computing, pp. 128–136 (August 2010)

28. Soh, L.K., Khandaker, N., Jiang, H.: I-MINDS: A Multiagent System for Intelligent Computer- Supported Collaborative Learning and Classroom Management. International Journal of Artificial Intelligence in Education 18(2), 119–151 (2008)
29. Strijbos, J.: Designing for interaction: Six steps to designing computer-supported group-based learning. Computers & Education 42(4), 403–424 (2004)
30. Tuckman, B.: Developmental sequence in small groups. Psychological Bulletin 63(6), 384–399 (1965)
31. Westera, W., Wagemans, L.: Help me! Online Learner Support through the Self-Organised Allocation of Peer Tutors. In: Abstracts of the 13th International Conference on Technology Supported Learning & Training, pp. 105–107. ICEW GmbH, Berlin (2007)

# Mastery Grids: An Open Source Social Educational Progress Visualization

Tomasz D. Loboda, Julio Guerra, Roya Hosseini, and Peter Brusilovsky

School of Information Sciences, University of Pittsburgh,
Pittsburgh, PA 15232, USA
{tol7,jdg60,roh38,peterb}@pitt.edu

**Abstract.** While many pieces of educational software used in the classroom have been found to positively affect learning, they often are underused by students. Open learning model and social visualization are two approaches which have been helpful in ameliorating that low usage problem. This article introduces a fusion of these two ideas in a form of social progress visualization. A classroom evaluation indicates that this combination may be effective in engaging students, guiding them to suitable content, and enabling faster content access.

## 1 Introduction

Over the last 30 years researchers and practitioners have developed a range of advanced educational tools for computer science education (CSE) such as animations, simulations, programming problems, and self-assessment questions. While a number of these tools have been evaluated through user studies and proven to be useful, a limited success of many others demonstrated that educational effectiveness observed in the lab does not necessarily translate into a broad educational impact [1]. One of the reasons for this seems to be low usage by students. That is, even in a favorable situation when a good tool is adopted by a motivated instructor and made easily available to students its use may be low. That plague of low usage afflicts advanced CSE tools the most because the majority of these tools are crafted to support knowledge enhancement through self-study which makes it hard to mandate and control their use. Consequently, unlike some assessment-focused tools which students are required to use in order to complete their assignments (e.g., program grading tools or program-construction questions), self-study tools are typically provided as a recommended content and, as such, their use rarely counts towards the final grade. Moreover, the motivation to work with self-study content might be decreased when multiple content items are available (e.g., many self-assessment questions or animated code examples) because then it is even harder to choose the most appropriate item to work on. Technologies aiming at increasing student motivation or guiding them to the most appropriate self-study content could help to ameliorate this low usage problem.

Past research has identified several such technologies. For example, *learner knowledge visualization* (known as *open learner modeling*) which acknowledges

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 235–248, 2014.
© Springer International Publishing Switzerland 2014

students progress and highlights gaps in their knowledge has been demonstrated to motivate student learning and guide students to the appropriate content [2]. The study of social visualization approaches based on the ideas of *social comparison* [3] has shown that the ability to compare oneself to one's peers or the entire class could increase a student's motivation to learn and improve their participation in educational activities [4,5].

The work presented in this article proposes a progress visualization approach that combines the prospects of open student modeling and social visualization in the form of social progress visualization. To that effect, we have developed Mastery Grids, a visualization tool that combines personal and comparative progress visualization. Mastery Grids has been designed as a tunable open source educational interface and can work with several kinds of self-study content and can be a part of semester-long classes. The remainder of this article presents the idea and the implementation of Mastery Grids and reports the results of its evaluation in two different domains and two kinds of self-study content.

## 2 Relevant Work

The idea of an open learner model was originally explored in the area of personalized learning systems. While in traditional personalized systems, learner models were hidden "under the hood" and used to make the education process personalized, the pioneers of open modeling argued that the ability to view and modify the state of one's knowledge could be beneficial for students. Open learner models for a set of skills or topics are most frequently displayed as a set of progress bars (also known as *skillometers*) that visualize learner knowledge state understood either as the probability that a learner has mastered a skill [6] or as a fraction of total knowledge to be gained for a topic [7] (which is sometimes estimated as a fraction of material covered by that topic). Other forms of knowledge visualization (e.g., graphical or color-coded) have been explored as well [8,9]. More recently, open learner models emerged as a feature of online learning systems which use some elements of progress modeling and personalization (e.g., Khan Academy Dashboard).

A range of benefits have been reported upon opening models to learners, such as increasing the learner's awareness of their developing knowledge, difficulties, and the learning process in general, as well as increasing students' engagement, motivation, and knowledge reflection [2,7,8]. Studies of both individual and group open learner models have shown an increase of reflection and helpful interactions among teammates. For more information, see Bull & Kay [2] who introduced a framework to apply open user models in adaptive learning environments and provided many in-depth examples.

## 3 Social Educational Progress Visualization (Mastery Grids)

Mastery Grid is out attempt to implement the ideas of social progress visualization in a generic, usable, and re-usable form. The most important component around which Mastery Grids is build is a *grid* which is a three-dimensional visualization component. The two geometric dimensions are topics defined in a domain (e.g., "Variables" or "Arrays"; the horizontal dimension) and resources which learners can access (e.g., questions, examples, readings, lecture recordings, etc.; the vertical dimension). The current classroom study used only two types of resources: Questions and examples. The third dimension is given by the intensity of the color in each cell of a grid and denotes the level of completeness, level of mastery, or level of progress a learner has for the given combination of topic and resource type. For example, the more saturated the color the more mastery the learner has achieved. This design allows learners to quickly get an idea about how they are doing in a course and where to allocate their efforts next.

Figure 1 shows the first of the two modes Mastery Grids can work in, the *All Resources* mode. In that mode, the "Me" grid shows progress the learner has made, the "Group" grid shows progress that the currently selected group has made, and the "Me vs group" grid shows the difference between the two grids. By manipulating the group the learner decides what they want to compare themselves to. For example, they could choose "class average" or "top 10 learners," both of which were used in the current classroom study. Below the three main grids are individual learner grids. Each learner in the currently selected group is represented by one grid and these grids are ordered according to the average progress.

In the All Resources mode described above a learner sees the entirety of the material which is available to them (i.e., all topics and all resource types). However, a learner can choose to focus on one resource in particular (e.g., questions) and enter the *Resource Focus* mode (Figure 2). In that mode, the three main grids are fused into a single main grid. That single grid displays only the selected resource rows from the three main grids, i.e., the first row shows the learner, the third row shows the group, and the middle row shows the difference between the two. The Resource Focus mode makes the me-group types of comparisons easier by displaying the relevant information closer together. As a result of focusing on one resource type, individual learners from the current group are shown in a single grid as well. A learner can control the height of rows of that grid to fit groups of even several hundred learners on one screen. The Resource Focus mode features a timeline which makes it easier for learners to identify which topic is currently being covered in class (the big circle). Furthermore, green and red circles respectively denote already-covered and to-be-covered topics. The timeline was not available to subjects in the current classroom study.

Each cell from either of the three main grids in the All Resources mode or the single main grid in the Resource Focus mode grants access to the list of activities. When a cell is clicked, topic grid (or grids) are replaced by activity grid (or grids) as shown in Figure 3. Each cell of an activity grid corresponds to a single activity the learner can access and clicking on that cell opens the activity window (Figure 4). To enhance the learning experience, when a learner answers a question incorrectly, a set of examples is recommended to them as remediation.

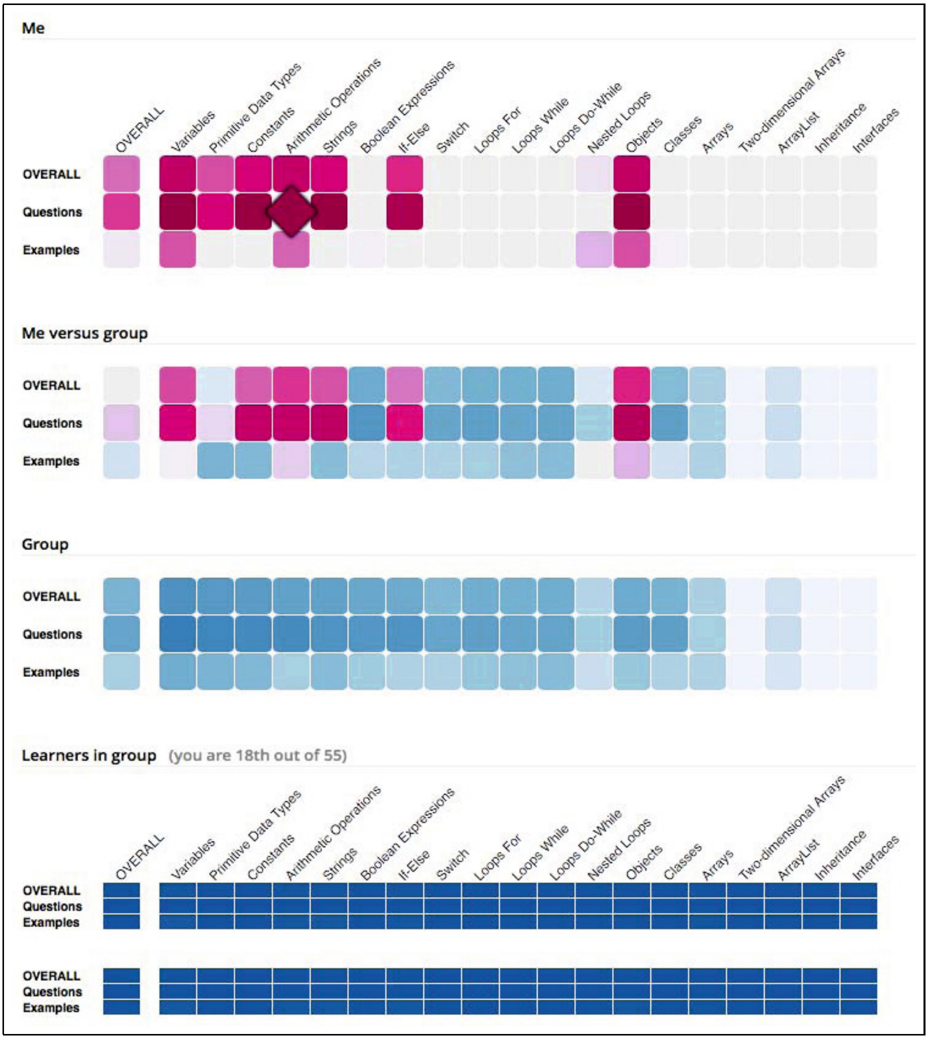



**Fig. 1.** Mastery Grids: All Resources mode. In this mode, the student sees all resources available to them. Each resource is a collection of activities of a different kind (here, Questions and Examples) and these resources are represented by rows of a grid. The columns of a grid show topic into which the course is divided. "Resources × topics" create a two-dimensional representation of the course material available through Mastery Grid. The third dimension is given by the intensity of the color of a grid cell; the more deep the color the more progress. The student can see their own progress on the first grid, the average progress of the current group (e.g., all students in the class) on the third grid, and the difference between the two on the second grid. A list of individual progress of all members of the current group is shown below and in the case of large groups the students would scroll down to inspect it (only two individual progress grids are shown here). The student's own progress is shown in that list as well so that they know where they are relative to other students. As indicated in the header of that section of the interface, the example student is 18th out of the total of 55 students.


**Fig. 2.** Mastery Grids: Resource Focus mode. Although many resources can be available to students, they may choose to focus on only one of them at any given point. For example, when they do not yet have a good understanding of the material, they may focus on readings or example source code snippets. Later on, however, especially before an upcoming in-class exam, they may choose to focus on evaluation-type resources to see how well they can answer questions related to the material. The Resource Focus mode gives students the ability to work with one resource only. The three main grids from the All Resources mode are summarized by only one grid. That grid shows the student's progress in the first row, the average progress of the current group (e.g., top 10 students in the class) in the third row, and the difference between the two in the second row. Below, a list of individual progress of all members of the current grid is shown. As shown on this screenshot, the student's own progress is displayed in that list as well; that gives them a visual clue about where they are relative to other students. This interface is more compact compared to the All Resources mode. Switching between the two modes happens through a simple always-on-the-screen toolbar which is not shown here.

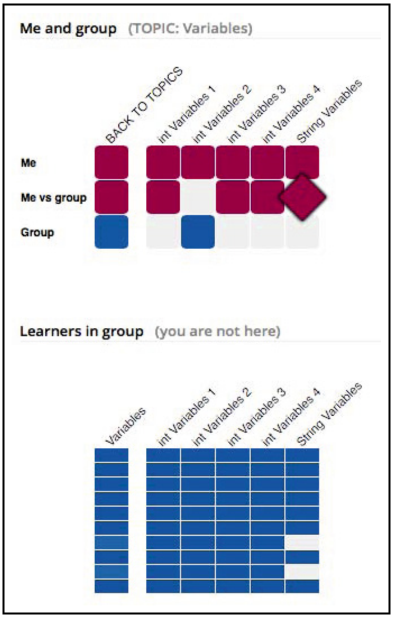



**Fig. 3.** Mastery Grids: List of questions for the Variables topic (Resource Focus mode). Irrespective of the mode the interface is in (i.e, All Resources of Resource Focus), when the student clicks on one of the cells of any of the main grids, a list of activities (see Figure 4) associated with that cell is shown. In this example, the interface is in the Resource Focus mode and therefore there is only one activities grid (vs. three for the All Resource mode) and it shows progress that has been made so far towards all activities for the selected resource (here, Questions) and topic (here, Variables). To invoke an activity, the student clicks on one of the cells of this activities grid.

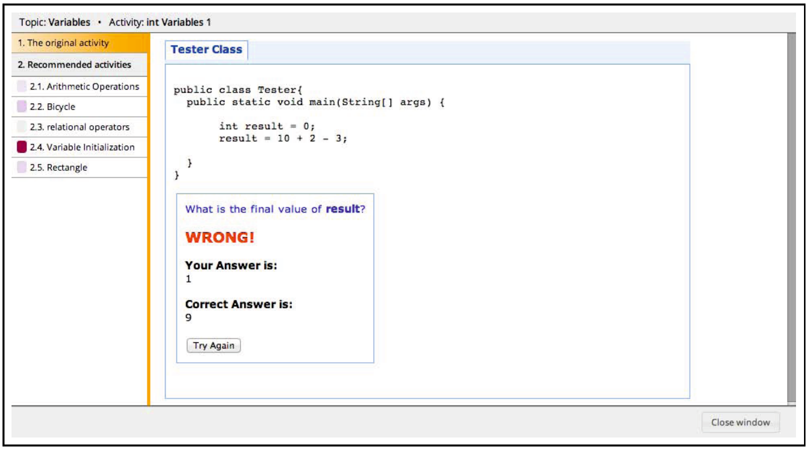



**Fig. 4.** Mastery Grids: A sample activity (here, a question) shown after the student clicks on one of the activity grid cells. In this example, the question has been answered incorrectly and as a result a list of recommended activities is shown on the left.

These examples are chosen based on the similarity between the question and all the examples available.

Technically, Mastery Grids is a Web application with the front-end written in HTML, CSS3, and JavaScript and developed by the first author. At the back-end, it is supported by Java servlets logic with a relational database persistence developed by the second and third authors. The front-end does not hard-code anything content-related and expects to receive everything from the server. In that respect, it is a flexible interface which can be deployed in many progress visualization scenarios.

## 4 Evaluation

### 4.1 Classroom Study

To examine the role of Mastery Grids as both a visualization and a content access interface we gave students an alternative way of accessing content through a simple two-level hierarchy of HTTP links which from now on we refer to as *Links interface* or simply *Links*. The first level links listed topics and the second level links listed activities (i.e., questions and examples).

A combination of Links and Mastery Grids was used in three courses in the Fall 2013 term at the School of Information Sciences, University of Pittsburgh. The three courses were: Introduction to Object Oriented Programming (undergraduate) and Database Management (graduate and undergraduate). Both tools were introduced to students by the first and second author on the second week of classes in the Java course and on the fourth week of classes in both database courses. Students were informed that the use of these tools was non-mandatory and that there was no penalty for not using them. To engage students, the Java course instructor offered extra points (5 out of 100) towards class participation for solving at least 15 questions using either Links or Mastery Grids[1]. Table 1 reports the number of students in each course along with the number of student sessions recorded.

**Table 1.** Classroom sizes and numbers of sessions in the three courses used in the current study

| Course | Students | Sessions | | |
|---|---|---|---|---|
| | | 0 | 1-3 | 4+ |
| O.O. Programming (Java) | 35 | 4 | 18 | 13 |
| Databases (grad) | 83 | 54 | 26 | 3 |
| Databases (undergrad) | 35 | 24 | 9 | 2 |
| TOTAL | 153 | 82 | 53 | 18 |

[1] None of the students did exactly 15 questions though.

### 4.2 Educational Material

The Links and Mastery Grids interfaces were used to provide access to and visualize progress in two kinds of self-study learning activities from two domains, Java and SQL. The first type of activities were examples delivered by WebEx system [10]. Each example was a snippet of Java or SQL code with annotated lines. All annotation were initially hidden and to see them a student needed to click on a line. The second type of activities were questions. Each question was a parameterized snippet of code that required students to mentally execute it and provide an answer. Java questions were delivered by QuizJet system [11] and SQL questions were delivered by SQL-Knot system [12]. Activities in both domains were organized into topics to reflect the structure of the three courses used in the current study. The Java corpus contained 19 topics with a total of 75 WebEx examples and 94 QuizJet questions while the SQL corpus contained 19 topics with a total of 64 WebEx examples and 46 SQL-Knot questions.

### 4.3 Usage Pattern Analysis

We used log analysis to understand how Links and Mastery Grids and their features were used by the students. Only Java class logs were used because they offered substantially more data than either of the other two courses[2].

Based on our analysis, students could be split into six groups based on how much they used both tools: (Z) Four students did not use either of the tools; (L1) Two students used Links only and never used visualization; (L2) Three students used Links for content access and loaded the visualization but did not use its interface; (L.MG1) 16 students used Links for content access and interactively explored visualization but did not use it for content access; (L.MG2) Six students used both Links and Mastery Grids for content access; (MG) Five students used exclusively Mastery Grids for content access. This data demonstrates a relatively high engagement levels in that while the use of Mastery Grids was not required, the majority of students used it as an information source and a considerable amount of them used it for content access despite the availability of a faster content access tool, Links.

To analyze the impact of Mastery Grids on student behavior, we combined the five groups of students who used either tool into two larger groups depending on their use of visualization (Table 2): Those who had little to no use of Mastery Grids (L1+L2; the two first columns) and those who used it considerably (L.MG1+L.MG2+MG; the two last columns). 26 students used Mastery Grids on about half of their sessions while only six did not use the visualization interactively. As the data shows further, students who used the visualization seemed to be more engaged with self-study content in that they answered more questions, tried more examples, inspected more example line comments, and got a higher correct question answer ratio. The groups did not differ with respect to

[2] We excluded student sessions shorter than 30 seconds. Additionally, we excluded two students which accessed a total number of questions larger than 3*SDs than the mean.

**Table 2.** Usage statistics by group (means; SDs in brackets)

| | L1+L2 | | L.MG1+L.MG2 | | MG | | L.MG1+L.MG2+MG | |
|---|---|---|---|---|---|---|---|---|
| Number of students in group | 5 | - | 20 | - | 6 | - | 26 | - |
| **Session info** | | | | | | | | |
| Number of sessions | 4.00 | (1.58) | 4.05 | (2.84) | 2.67 | (2.07) | 3.73 | (2.71) |
| Number of MG sessions | 0.80 | (0.84) | 3.00 | (2.00) | 2.67 | (2.07) | 2.92 | (1.98) |
| Number of MG UI actions | 0.00 | (0.00) | 25.30 | (57.48) | 111.83 | (82.99) | 45.27 | (72.60) |
| Sessions with 1+ MG UI actions | 0.00 | (0.00) | 0.46 | (0.50) | 0.94 | (0.25) | 0.54 | (0.50) |
| Sessions with 5+ MG UI actions | 0.00 | (0.00) | 0.14 | (0.34) | 0.69 | (0.48) | 0.23 | (0.42) |
| Sessions with 10+ MG UI actions | 0.00 | (0.00) | 0.10 | (0.30) | 0.62 | (0.50) | 0.19 | (0.39) |
| Duration (h:mm) | 1:00 | (1:46) | 1:13 | (1:53) | 0:24 | (0:30) | 1:05 | (1:46) |
| **Question answering performance** | | | | | | | | |
| Correct answer ratio | 0.58 | (0.11) | 0.69 | (0.13) | 0.73 | (0.12) | 0.70 | (0.13) |
| **Material exposition (total number of accesses)** | | | | | | | | |
| Activities | 86.60 | (53.94) | 111.80 | (96.70) | 68.17 | (46.41) | 101.73 | (88.82) |
| Questions | 66.20 | (37.37) | 85.65 | (81.83) | 61.00 | (44.14) | 79.96 | (74.77) |
| Examples | 20.40 | (20.51) | 26.15 | (28.02) | 7.17 | (6.94) | 21.77 | (25.94) |
| Unique activities | 58.00 | (33.90) | 69.75 | (51.90) | 45.00 | (26.84) | 64.04 | (48.00) |
| Unique questions | 39.80 | (22.10) | 45.30 | (29.54) | 38.00 | (23.19) | 43.62 | (27.94) |
| Unique examples | 18.20 | (15.79) | 24.45 | (25.43) | 7.00 | (6.63) | 20.42 | (23.59) |
| Example lines | 55.00 | (36.10) | 106.70 | (133.99) | 29.67 | (44.44) | 88.92 | (123.02) |
| Unique example lines | 47.40 | (30.28) | 93.75 | (115.42) | 22.33 | (32.30) | 77.27 | (106.19) |
| **Material exposition (rate, i.e., total number of accesses per minute)** | | | | | | | | |
| Activities | 0.84 | (0.82) | 0.66 | (0.47) | 1.27 | (0.61) | 0.80 | (0.56) |
| Questions | 0.71 | (0.74) | 0.53 | (0.43) | 1.15 | (0.66) | 0.68 | (0.55) |
| Examples | 0.13 | (0.09) | 0.12 | (0.09) | 0.12 | (0.15) | 0.12 | (0.10) |
| Unique activities | 0.58 | (0.58) | 0.44 | (0.30) | 0.89 | (0.40) | 0.54 | (0.37) |
| Unique questions | 0.45 | (0.49) | 0.32 | (0.26) | 0.77 | (0.43) | 0.43 | (0.35) |
| Unique examples | 0.13 | (0.10) | 0.12 | (0.08) | 0.12 | (0.15) | 0.12 | (0.10) |

how quickly they explored the material space (measured by activity access per minute).

To examine the difference between students who used Mastery Grids exclusively and those who used it only some of the time, we also contrasted the MG group with groups L.MG1 and L.MG2 combined. According to data reported in Table 2 (3rd through 8th columns of numbers), students from the MG group worked with the content more productively by accessing questions (both those they had already seen and those they had not) at a higher rate than students from the two other groups. We return to this observation in the discussion.

While these results seem encouraging, they are based on a small number of students which incidentally prevented us from performing statistical inferences. Because of that, the above analysis needs another evaluation done in light of better data. In the future, we hope to address this concern and thus get a better insight into the different use patterns.

### 4.4 Material Exposition and Final Grades

While data from the current study cannot be used to draw causal links between material exposition and final grade (or any other indicators of educational performance) we wanted to discover if more student activity was associated with worse or better final grade. To that effect, we fitted four linear mixed models, each having the final grade as the response variable. To use all data we had at our disposal, we pooled data from all three courses described earlier in this section. To properly model the structure of the combined data, we included class random effect as a higher level unit. Student random effect was nested in the class random effect. Because relatively few students from the two database courses used the supplementary tools at all, we considered only students who logged in at least once (we required no activity beyond that though).[3]

First, we wanted to see if more interaction with the material as a whole could have affected the grade. To do that, we checked if the total number of questions and the total number of examples accessed (through folders and visualization together) were significant predictors of the final grade. After discovering that they were not, we broke the material apart into two parts and asked if perhaps the amount of content accessed only through folders or only through the visualization had a significant association with the grade. That, however, was not the case either. To summarize, we did not establish any link between the shear amount of interaction with the material and the grade.

Because the visualization offers an elaborated user interface that presents students with a lot of information, we decided to broaden our search and see if increased student activity in the visualization itself could be held responsible for a change in the grade. In order to address this question, we looked at the number of *educational actions* students performed while using Mastery Grids. That number was a sum of the number of times students performed any of the following eight actions: (1) Selecting a topic; (2) Accessing a question;

[3] Using all student data did not change the results substantially.

(3) Accessing an example; (4) Accessing a recommended example; (5) Accessing the original question again after inspecting at least one recommended example; (6) Clicking on a line of an example to reveal the comment associated with that line; (7) Assessing difficulty of a question; (8) Navigating back to topics grid from activity grids. We found a statistically significant amount of support for this relationship. More specifically, performing one educational action was associated with an increase of 0.016 in the grade (SE=0.007; $p$=.0187). That is, performing a 100 such actions would be associated with a 1.6 increase in the grade. If such an increase was to occur, then that would be a reasonable effect size.

### 4.5 Subjective Responses

To understand how students perceived Mastery Grids, we asked them to complete a survey consisting of two parts:[4] (1) general evaluation containing 14 questions addressing usefulness, usability, and satisfaction, and (2) feature-specific evaluation containing 15 question about the most important components of Mastery Grids.[5] All questions were phrased as positive statements and a 5-points Likert scale was used to collect students' responses (*strongly disagree*, *disagree*, *no opinion*, *agree*, and *strongly agree*).

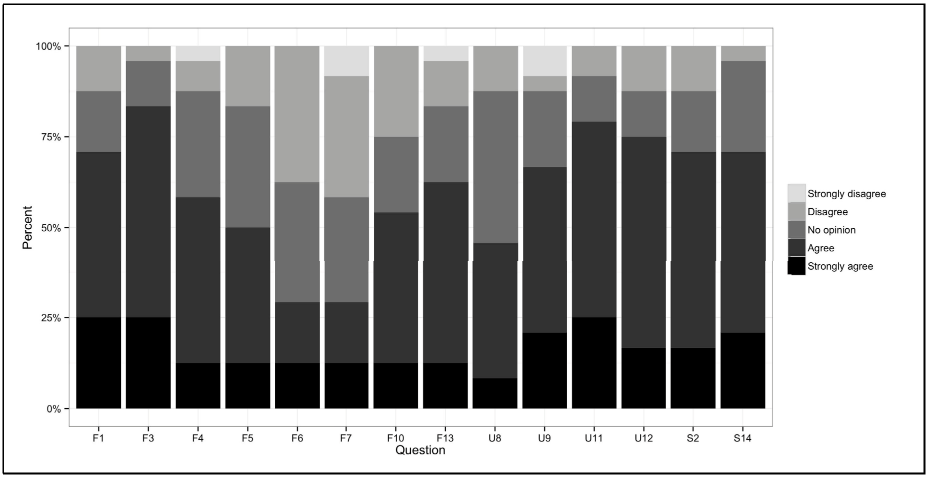



**Fig. 5.** Subjective responses: General (F–Usefulness; U–Usability; S–Satisfaction)

[4] Due to space limitations we do not provide text of questions but we refer to them as much as possible in the main text.

[5] Screenshots were used to indicate the parts of user interface questions were referring to.

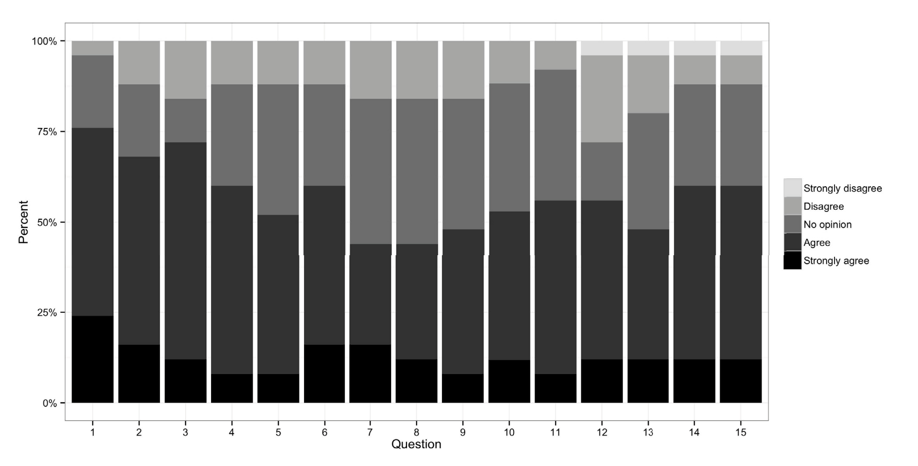



**Fig. 6.** Subjective responses: Feature-specific

A total of 39 students from the three courses completed the survey. However, we excluded three students who answered all questions giving the same score, and 11 other students who used either Links or Mastery Grids in no more than two sessions each shorter than 30 seconds. Thus, responses of 25 students were considered in the analysis.

The survey data shows that students' perception of Mastery Grid's usefulness and usability was positive (Figure 5). The majority of the students agreed or strongly agreed that the application was useful (F1, 70.8%), that they liked the tool and would like to use it in other courses (S2, S14, 83.3% and 70.8%, respectively), that the tool was easy to use (F10, 66.7%), and that it was easy to remember how to use it (U11, 79.2%). Interestingly, social comparison aspects of progress visualization were perceived more positively than personal aspects. That is, while the majority of students (F7, 54.2%) agreed or strongly agreed that viewing classmates' progress motivated their own progress, less than a third agreed or strongly agreed that progress visualization helped to identify their weak points (F5, 29%) or plan their class work (F6, 29%).

The feature-specific responses (Figure 6) were also generally positive: 76% and 68% of the students agreed or strongly agreed that the "Me" grid was useful for viewing one's own progress at the level of topics and activities, respectively (questions 1 and 2). Similarly, 72% and 60% of students agreed or strongly agreed that the "Group" grid was useful at the level of topics and activities, respectively (questions 3 and 4). However, a smaller fraction of students thought that it was useful to be able to compare oneself against the group using either the "Group" grid (52%) or the "Me vs group" grid (44%). The usefulness of the individual students' progress (the grids displayed below the main grids) was evaluated positively by about half of the students (48% for topics, question 9; 52.9% for activities, question 10).

Regarding the feature-level usability, 60% of students thought that progress at both the topic level and the activities levels was clearly shown. However, only 48% of the students stated that they quickly learned how to go from topics to activities and back (question 13) which could explain why majority of students used Mastery Grids interface to view their progress and progress of others and only a third of them used it for content access.

## 5 Discussion and Conclusions

In this article, we have presented Mastery Grids, an open source implementation of a social progress visualization interface. We have also reported results of a classroom study that examined how students used the system as well as their feedback. In that semester-long study, students in three courses were offered access to two kinds of advanced self-study content and progress visualization through Mastery Grids. They were also offered a simpler and more traditional way of accessing content we have called Links.

When analyzing usage patterns we observed that students who used Mastery Grids had a higher ratio of questions answered correctly than those who used Links only. It is possible then that the visualization guided students to questions which were more suited for their level of understanding of the material. We also observed that the visualization directed students to new material at rates higher than the alternative Links interface. A similar finding has been reported before in the context of adaptive explanatory visualization [13]. The mechanisms responsible for elevating material exposition rates in both the case of that earlier report and the current study are likely the same and may be the result of the visualization attempting to stay in sync with students' progress and thus being able to direct them to new content more quickly. If that was indeed the case, then our results would underline the importance of accurate student modeling.

Our finding of the number of educational actions performed in Mastery Grids being predictive of the final grade is a hopeful one. After all, a tool like that should help students to climb the mountain of scholastic achievements. However, because of the lack of manipulation, we recommend that this relationship be studied more systematically. More importantly, the current study cannot adjudicate if it is indeed the case that using the visualization more helped students with getting a better grade or if instead students which ended up getting a better grade were also the ones more likely to be engaged with supplementary educational tools.

Finally, student feedback analysis demonstrated that students assessed the usefulness and usability of the system quite positively. At the same time, some features of the system were regarded as less positive than others indicating that there is room for improvement for the future versions of Mastery Grids.

## References


1. Naps, T., Rößling, G., Anderson, J., Cooper, S., Dann, W., Fleischer, R., Koldehofe, B., Korhonen, A., Kuittinen, M., Leska, C., McNally, M., Malmi, J., Rantakokko, J., Ross, R.J.: Evaluating the educational impact of visualization. ACM SIGCSE Bulletin 35(4), 124–136 (2003)
2. Bull, S., Kay, J.: Student Models That Invite the Learner In: The SMILI() Open Learner Modeling Framework. International Journal of Artificial Intelligence in Education 17(2), 89–120 (2007)
3. Festinger, L.A.: Theory of social comparison processes. Human Relations 7(2), 117–140 (1954)
4. Linton, F., Schaefer, H.-P.: Recommender systems for learning: Building user and expert models through long-term observation of application use. User Modeling and User-Adapted Interaction 10(2-3), 181–208 (2000)
5. Vassileva, J., Sun, L.: Using community visualization to stimulate participation in online communities. e-Service Journal 6(1), 3–39 (2007)
6. Corbett, A.T., Anderson, J.R.: Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction 4(4), 253–278 (1995)
7. Mitrovic, A., Martin, B.: Evaluating the effect of open student models on self-assessment. International Journal of Artificial Intelligence in Education 17(2), 121–144 (2007)
8. Zapata-Rivera, J.-D., Greer, J.E.: Interacting with inspectable Bayesian student models. International Journal of Artificial Intelligence in Education 14(1), 1–37 (2004)
9. Hsiao., I.H., Bakalov, F., Brusilovsky, P., König-Ries, B.: Progressor: Social navigation support through open social student modeling. New Review of Hypermedia and Multimedia 19(2), 112–131 (2013)
10. Brusilovsky, P., Yudelson, M.V., Hsiao, I.H.: Problem solving examples as first class objects in educational digital libraries: Three obstacles to overcome. Journal of Educational Multimedia and Hypermedia 18(3), 267–288 (2009)
11. Hsiao, I.H., Sosnovsky, S., Brusilovsky, P.: Guiding students to the right questions: Adaptive navigation support in an e-learning system for Java programming. Journal of Computer Assisted Learning 26(4), 270–283 (2010)
12. Brusilovsky, P., Sosnovsky, S., Lee, D.H., Yudelson, M.V., Zadorozhny, V., Zhou, X.: An open integrated exploratorium for database courses. In: Proceedings of 13th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE), pp. 22–26 (2008)
13. Loboda, T.D., Brusilovsky, P.: Adaptation in the context of explanatory visualization. In: Dillenbourg, P., Specht, M. (eds.) EC-TEL 2008. LNCS, vol. 5192, pp. 250–261. Springer, Heidelberg (2008)

# Abstraction of Learning Management Systems Instructional Design Semantics: A Meta-modeling Approach Applied to the Moodle Case-Study

Esteban Loiseau, Pierre Laforcade, and Sebastien Iksal

LIUM (Computer Science laboratory)
University of Le Mans, France
{esteban.loiseau,pierre.laforcade,sebastien.iksal}@univ-lemans.fr
http://www-lium.univ-lemans.fr

**Abstract.** Nowadays Learning Management Systems (LMS) are not restricted to distant uses. Nevertheless, the pedagogical expressiveness of designed courses is strongly dependent on the teacher's knowledge and expertise about the targeted platform. The funded GraphiT project aims to help teachers in focusing on the specification of pedagogically sound and technically executable learning designs. To this end, we propose to support teachers by providing an LMS-specific Visual Instructional Design Language (VIDL). This paper proposes a specific LMS- centered approach for raising the pedagogical expressiveness of its implicit learning design semantics. We discussed, in accordance to the activity theory, how the LMS low-level parameterizations could be abstracted in order to build higher- level building blocks. Based on the Moodle application, we present and illustrate our approach by formalising the abstract syntax of a Moodle-dedicated VIDL.

**Keywords:** Instructional Design, Visual Instructional Design Language, Learning Management System, Modeling and Meta-modeling.

## 1 Introduction

Nowadays, *Learning Management Systems* (LMS) are widely spread in academic institutions. They are not restricted to online and distant courses but are also useful during or in addition to face-to-face learning sessions [11]. Nevertheless, the results of a study we conducted with 203 teachers, put forward their heavy form-oriented human-interfaces and tools/services-oriented course design lead to reduce their uses. In order to set up complex learning activities, teachers must develop high-level skills on how to use the existing LMS: how and when managing and sequencing the available features and tools for pedagogical purposes. Such skills can be acquired through specific teacher education programs, often focusing on the features and technical aspects of the platform, but few courses deal with how to design pedagogically sound learning situations on the LMS

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 249–262, 2014.
© Springer International Publishing Switzerland 2014

(specific learning design). Because of the multiple educational theories [16] and approaches, as well as the lack of tools and processes dedicated to existing LMSs, teachers develop *ad hoc* and individual learning design techniques.

In such context, it seems relevant to help teachers in focusing on pedagogical aspects and their instructional design setting-up for the specific LMS they have at their disposal. Whereas improving their know and know-how about the platforms features, a focus on the instructional design possibilities and how they can rely on the platform features should encourage individual and collective understanding about the pedagogical uses of the targeted LMS.

We on purpose propose an LMS-centered designing approach in opposition to the usual platform-independent approaches [3,13]. The GraphiT project (funded by the French Research Agency) is based on this approach. Its main objective is to investigate Model Driven Engineering (MDE) and Domain Specific Modeling (DSM) techniques to help specifying LMS-centered graphical instructional design languages and development of dedicated editors. This paper focuses on the main challenge: raising the pedagogical expressiveness of the LMS learning design semantics by using meta-modeling techniques.

To this end, we detail in Section 2 our research context, including the presentation of the GraphiT project from a MDE and DSM perspective. The section 3 is dedicated to the presentation of our abstraction approach. We analyse users' activities on an online course according to the activity theory. We also detail the designers requirements and needs collected during interviews. We also formalise, as a metamodel, the concrete Moodle application of our approach. We explain and illustrate the 4-levels architecture we propose in Section 4 and illustrate it by a simple but representative example in Section 5. We also illustrate, in Section 5, some concrete mapping examples we tooled by using some specific DSM tools. The final conclusion sketches the next steps we are tackling to drive the elaboration of the graphical instructional design language from this abstract syntax.

## 2 Research Context

### 2.1 LMS and Instructional Design

LMSs development is usually based upon an educative theory rationale, or some specific pedagogical approach. For example Moodle claims a socio-constructivist pedagogy philosophy [9]. Most spread LMSs generally follow such an orientation because of the various production and communication tools provided. LMSs are the activity-centered evolution of former learning-objects-centered TEL-systems. Indeed, current LMSs provide designers with some numerous functionalities that can be used to rrealisevarious learning activities and are not restricted to give resources access to individuals.

Nevertheless, activity-centered standards like the *de facto* IMS-LD fail to integrate existing LMSs. Experiments on extending Moodle to import IMS-LD learning scenarios proved that adapting existing LMSs requires some complex and heavy re-engineering (in particular integrating a dedicated runtime-engine) in order to overcome the limits of the platform features and semantics [7]. Educational Modeling Languages [5] fail to provide a support for operationalizing EML-conformed learning scenarios into existing LMSs. For now widely spread LMSs like Moodle still do not propose an IMS-LD compliance but a SCORM one if administrators decide to enable such import abilities to the LMS instances available to teachers-designers.

Moodle proposes its own format for importing questions into quizzes. Our idea is to generalise it to the whole instructional design aspects. Similarly to the SCORM compliance about Learning Objects, the rationale of the GraphiT project is based on the idea that LMSs should make explicit their learning design format in order to ease the import/export of learning scenarios.

### 2.2 Overview of the GraphiT Project from an MDE and DSM Perspective

The project main goal is to study the possibilities and limits about the pedagogical expressiveness of operationalizable languages. The project methodology consists in exploring how *Model Driven Engineering* and more particularly *Domain Specific Modeling* techniques and tools can be relevant and useful to achieve this goal.

Similar research works follow different approaches. For example the *Glue! architecture*, including the *Glue!PS* editor [3], and the CADMOS editor [13] are LMS-independent solutions offering an LMS deployment feature towards the most spread and used Moodle platform [15].They both achieve the deployment by generating a Moodle course backup with all the information, mapping their own data model concepts to Moodle data model concepts; this backup is then imported and deployed within a Moodle course using the Moodle restoration process. Such approaches result in semantics adaptations and semantics losses during their internal mappings because of the gap between the instructional design language and the specific learning design capabilities and features of the targeted LMS. Other works [1] shows that transformation models techniques from the MDE theories and tools can be useful to translate a designer-centered and LMS-independent learning scenario to a specific LMS one. Nevertheless, they also highlighted the complex transformation model to specify, the LMS metamodel to capture, the semantics losses during translation, and the requirement of an LMS- dedicated tool for embedding the scenarios into the LMS.

Our approach is different: we propose an LMS-dependent architecture that only focuses on one existing LMS in order to provide an instructional design language that will be specified and tooled according to the future mappings to realise. Our idea is to conduct the platform abstraction in accordance with the formalisation of future learning scenarios. We do not aim at extending the LMS semantics with new add-ons/plugins enriching it with more pedagogical-oriented features. Our objective is to support learning scenarios specification in conformance with the LMS semantics (its abilities as well as its limits). We also do not aim at only providing a notation layer on top of the LMS metamodel. By extending the LMS metamodel we also extend the abstract syntax of the instructional design language and then losing the LMS-compliance format. We plan to restore it by DSM techniques (weaving and transformation models) we are currently experimenting. We aim at guarantying that learning scenarios could be fully operationalized into the LMS without semantics losses. Obviously, our approach can take advantage of this LMS-dependance but it has also the inconvenience to be restricted to one LMS and one of its versions.

A global architecture of our solution is illustrated in Figure 1. The LMS instructional design semantics has first to be identified and formalized as a domain metamodel. This metamodel drives the elaboration of an XSD schema that will be used as a format reference for the API to develop. This API will be used through an import facility available to teachers-designers in their LMS courses. It will take in charge the XML-based scenario parsing and the LMS's databases filling-up. The LMS metamodel will also act as a basis for the elaboration of the visual instructional design language. According to DSM techniques and tools (like the EMF/GMF ones for example), this language will be composed of an abstract syntax from which the graphical, tooling and mapping models will be derived. The editor will also be developed using the code-generation feature of DSM tools.

The produced scenarios have to be compliant with the initial LMS meta-model to be deployed by the API. We propose then to run two models transformations. The first one will consist of various, fine-grained transformations during design-time: it will show some LMS mappings to teachers-designers in order to help and guide them in the design process. They are endogenous transformations because source and target models will be both conformed to the instructional design metamodel. The second transformation, unique and large, will be used as an export feature (after design-time). This exogenous transformation will produce a scenario/model conformed to the LMS-metamodel.

Past works have focused on the LMS meta-model formalization [2]. We are currently focusing on the abstract syntax of the instructional design language.

### 2.3 Focus on the Instructional Design Abstract Syntax from a Metamodeling Perspective

The main challenge of this project is to abstract enough the LMS instructional design semantics to provide teachers with some pedagogically-sound higher design blocks. The LMS expressiveness and limits have to be overcome in order to offer teachers some instructional design mechanisms closer to their practices and needs about specifying and sequencing learning activities. Concretely, the issue is about how specifying the metamodel of the instructional design language (MM-ID) in close relation to the LMS one (MM-LMS).

To this end, we already led some experiments [14] about three different approaches: 1/ MM-ID and MM-LMS are two different metamodels without any structural relations, 2/ MM-ID and MM-LMS are the same, the ID-language is only built upon the graphical concrete syntax, 3/ MM-ID is an extension of MM-LMS. The first approach corresponds to the usual way to specify an instructional design language with its main advantage (expressiveness) but also inconvenience

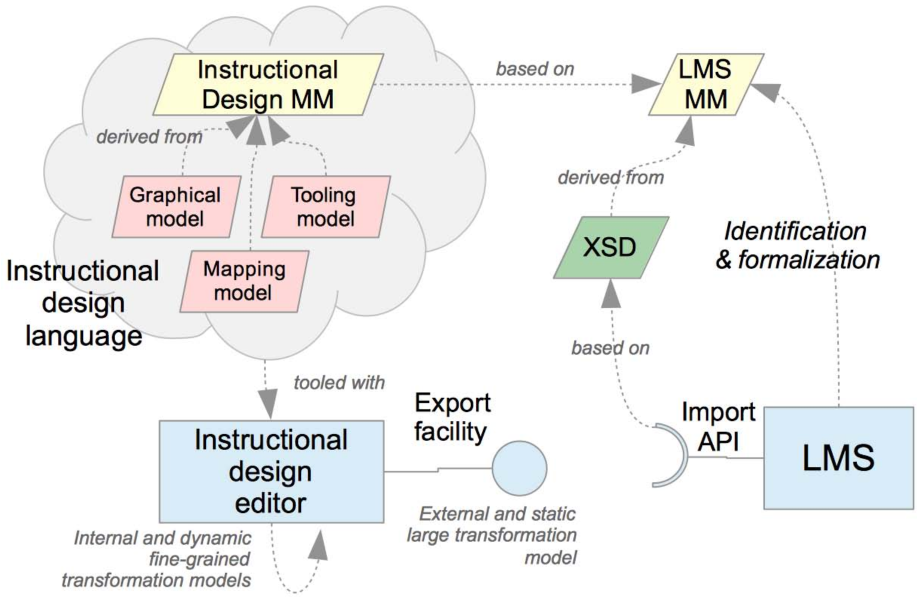



**Fig. 1.** Global overview of the GraphiT architecture

(difficulty to operationalize). The second one reveals the limits of the concrete syntax expressiveness when only defined by derivation of the abstract syntax. Finally, the third approach is intermediate on all criteria: best expressiveness / LMS compliance ratio. However, it requires a strong metamodeling expertise to reduce the developing cost while restoring the LMS compliance. This approach also highlights the importance to drive the expressiveness (and semantics) extension of the initial metamodel with the binding capacity. This paper focuses on our further results and propositions about this issue.

### 2.4 Analyzing LMS Uses from an Activity Theory perspective

According to the activity theory [6], the activity is the minimally meaningful unit of study to consider. We can then analyze the activity of learners/tutors from the activity theory perspective. By analogy we analysed the minimal actions one can attempt to realise when accessing an LMS course. We exemplify this analysis through specific Moodle cases.

The *subject* refers to an individual (learner/tutor) or a group of individuals who are enrolled and involved in the course. The Moodle group concept is out of our scope because of our generic design approach (a scenario should be run several times for different enrolled students). The *object* refers to material/concepts, like resources, to be transformed into *outcomes* according to the course objectives. The *instrument* can also be any Moodle tools (Moodle refers to them as activities) used to help the transformation process. The *community* is a group of learners/tutors who share the same general object. In Moodle, role refers to a collection of permissions defined for the whole system. Although they can be managed by the LMS administrators they are only useful for pedagogical purposes when learners have to participate in the design of the course themselves. Otherwise the grouping concept allows restricting specific activities and resources to specific groups of learners. It is then pedagogically interesting to declare them as communities when designing a learning scenario. Their uses also correspond to the *rules*, governing behaviors of community members within activities, as well as the *division of labor*, that is the distribution of tasks and powers between the members of the community. This division of labor relies on several Moodle concepts for sequencing (section, elements indentations (horizontal position within the section), restrict accesses, achievement conditions) and showing (visible/non-visible, restrict access information display) of courses elements.

This analysis highlights the idea that LMSs support the setting-up of activities according to the different components of the activity theory. It could be useful to help and guide teachers to think more about the design of their courses in terms of more abstract pedagogically-oriented activities instead of focusing on the technical-oriented setting-up of activities. To this end, it is relevant to consider the LMS instructional design metamodel as a basis from which an higher abstraction level could be raised.

### 2.5 A Practical Overview of Teachers' Requirements

Before tackling the LMS metamodel extension we first had to collect LMS-specific pedagogical needs and practices. We then conducted several theoretical, from literature sources [8], and practical studies with surveys and iterative interviews of 203 teachers and pedagogical engineers. These interviews covered a large variety of LMS use contexts: online learning, local learning within universities, full distant courses as well as blended learning. Although the GraphiT project deals with different LMSs for guarantying the reproducibility of its results, we on purpose propose to focus on the Moodle platform which is the most popular open-source Learning Management System. The analysis of these different sources aimed at collecting pedagogical practices or needs, and requirements about the languages and editors to specify and develop.

Most regrettable point highlighted is that practitioners do not really have complex practices to capture, because of the heterogeneity of their Moodle expertises and pedagogical backgrounds. Nevertheless they have in common to think about Moodle tools according to their basic pedagogical uses. Indeed, they all point the heavy parameterizations of tools and resources and the need for having an abstract view of what are the pedagogical uses in order to help and guide them in selecting and configuring the right implementation activities.

From advanced studies with pedagogical engineers we listed some specific requirements for language/editor to develop. First, they mentioned the need for the graphical authoring-tool to allow designers to select pedagogical blocks on top of the LMS semantics as well as with Moodle building blocks to compose with. In their mind, the editor will not have to strictly follow a top-down process from abstracted specification elements to implementation one expressed in terms of Moodle; abstractions from Moodle and its own concepts should be mixed up together according to practitioners' expertise about instructional design (**specification and implementation concepts mix**). Secondly, they are interesting in the idea that mappings from pedagogical design blocks to Moodle concepts can be showed to practitioners (**default mapping**) and adapted if required (**mapping adaptation**). This design approach could help practitioners in the appropriation of the pedagogical constructs and guide them in designing more abstract learning scenario while mastering the translations into LMS elements.

Another design point highlighted (**declarative non-visible information**) is about the possibility to design and declare within the learning scenario some information that do not required to be mapped into LMS concepts or just mentioned as non-visible labels (for students/tutors) for the teacher him-self: information about the face-to-face sessions mixed up with the LMS-centered ones, about pedagogical strategies or pedagogical objectives, about activities to realize on the LMS at a specific runtime moment according to concrete data (enrolled students, dates, etc.). Finally, another design need was to help teachers in sequencing the course in more advanced structures (choices, sequences with elements showed one-by-one according to their progress (**advanced activity structures**). Indeed, these can be done manually but it requires to parameterize many low-levels and technical-oriented properties (achievements, restricted access conditions...) that they would appreciate not to have to set up by themselves.

## 3 Formalization by Metamodeling

According to these practitioners' needs we propose to drive the abstraction of LMSs semantics by raising the LMS uses supporting learners/tutors activities. The following sections present these abstractions in relation with their formalizations for the Moodle LMS (Figure 2). We used the Ecore metamodel format because it will be handled by the EMF and GMF metamodeling tools [10] in order to drive the specification of the instructional design language and the development of its dedicated graphical tool. The metamodel from Figure 2 can be considered as the general abstract syntax of the instructional design language to be developed.

### 3.1 Fine-grained Pedagogical Activities as First Abstraction

The first LMS-abstract building block we propose is the pedagogical activity. We define this activity as an *abstraction of parameterizations one can realise when using a LMS tool or resource for a specific pedagogical usage.* From a single tool, for example a forum, one can design several pedagogical uses, depending on its configuration: to provide news to students, to set up group work, to propose a peer reviewing activity, etc.

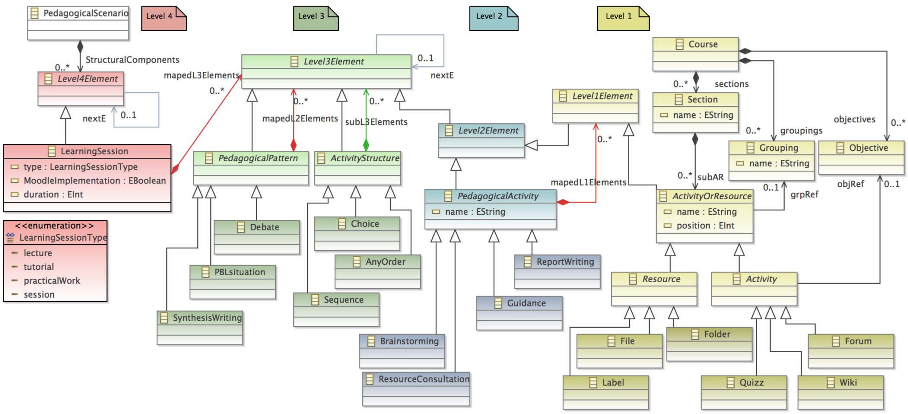



**Fig. 2.** The 4-levels abstract syntax of an instructional design language on top of the Moodle metamodel

Because several LMS functionalities can be used for the same pedagogical purpose, we have to find the discriminatory criteria that can guide to identify the right tool and default configuration (as well as the relations to objectives, resources, groupings, etc. that are involved in the right setting-up of the pedagogical activities).

To be used appropriately, this first abstract block requires a name, a description, and specific properties (the former discriminants), set at design-time by practitioners, that will drive the default mapping. For example an exchange activity, involving student communication, could either rely on a forum or a chat, depending on a synchronous property. The mappings will not be limited to the parameterization of a tool. For example, it will also impact some other elements in relation with the tool/resource: grades, objectives, groupings, restriction access and achievements rules, etc.

### 3.2 Large-grained Pedagogical Activities as Second Abstraction

The second LMS-abstract building blocks are of two kinds. We propose to adapt and integrate some pedagogical patterns and templates from literature [4,12] for examples as high-level blocks to use and combine for building learning sessions involving instructional strategies: inquiry, problem solving, role-playing, exploration, etc. Although practitioners from our studies do not use to compose with them, we aim at integrating them to encourage some pedagogical reflections and to guide designers towards new ways to realise their didactic and pedagogical objectives. This kind of pedagogical patterns will also have a description of their context, problem and solution uses. They will rely on a mix of structural activities, low-levels blocks (pedagogical activities) and LMS elements.

In order to ease and assist the practitioners when assembling and setting-up combinations of activities or resources we propose usual structural elements (selection, sequence, conditional activities, etc.). These blocks will be composed of other blocks, from high or low levels, including themselves. Every instructional design language feature some of them. In the case of Moodle they will be concretely translated as complex combinations of labels (stating the structure kind and use for users) and shifted content (move left/right Moodle feature) according to the activity structure components in the learning scenario. After various translations and mappings until reaching the LMS low-level elements, all its content parts will be parameterized (restrict access, visibility, achievement...) with appropriate properties in order to set up the desired behavior.

### 3.3 A 4-Levels Abstract Syntax

The global architecture we propose for the abstract syntax of the Moodle-centered instructional design language is composed of four levels. Figure 2 illustrates our proposition with a graphical representation of the ecore domain model.

Level 1 fits the Moodle metamodel. Readers have to consider Figure 2 as an incomplete representation of the whole metamodel. Only important structural relations and concepts are depicted. *Level 1 elements* (restricted to the Moodle *activities* the Moodle name given to the tools - and resources) can be directly used by teachers-designers and parameterized for building a learning session. From the Moodle metamodel point of view these elements require a global *Course* and a *Section* container to be attached to. In the extended metamodel they will be specified at first as child of *level 4 elements*. The model transformation, at post-design-time, will deal about restoring a model in full-compliance with the Moodle metamodel: creation of the global *Course* instance, Section instances, attachment of all the corresponding Moodle elements according to the orders and positions deductible from the source scenario.

Level 2 include our first high-level blocks about pedagogical activities. They are composed of *Level 1 elements*, i.e. Moodle activities and resources. Level 3 captures the second abstract blocks about pedagogical patterns and activity structures. The first one will be composed, after the design-time transformation model, by *Level 3 elements* that includes those from levels 1-to-3 including structural activities and other *Pedagogical Activities*. The activity structures are also composed of *Level 3 elements* but their content will be specified during the design-time. Finally, the fourth level is the contextual level focusing on the global structure of the learning session in relation to the different face-to-face, complementary or distant sessions.

Such *Level 4 elements* rely on the Moodle *section* concept. Indeed, Moodle only proposes sections into the space of the course for aggregating the tools and resources. However, designers have at their disposal an *indentation* feature (*position* property in the Moodle metamodel) to shift activities and resources in order to visually indicate their collective relationships. This *position* property will be used by the dynamical mappings, in order to position the corresponding elements in accordance to the source element position in the global learning scenario. The relations with a red composition indicate that the content will not be showed in the future concrete syntax (notation) as nested elements but will be shown in another sub-diagram where the parent container will play the role of the root canvas. Differently, the green composition indicates that content will be showed as nested elements of the parent container in the same diagram. Finally, the *nextE* reflexive relation allows, by inheritance, to provide a previous/next information to sequence the various elements within their dynamic pedagogical context (the ordering concerns the child elements sharing a same *Level Element* parent). The future authoring-tool will directly propose to practitioners the *level-4 elements* in the tool palette. Indeed, these elements are necessary to map to Moodle sections in order to sequentially structure the course skeleton. Sessions that do not rely on Moodle features can also be described if designers need an overall view of a global module/course larger than the ones involving the use of an LMS. Other *level-4 elements* will then open an empty sub-diagram when double-clicked. It can then be used to arrange *levels 3-to-1 elements* from the new palette. Indeed, practitioners can then choose the method (top-bottom, bottom-up), the description level (specification versus implementation) and the elements to select, combine and adapt. Except activity structures, other *levels 3-to-2 elements* can be opened up as another sub-diagram containing the default mapping to *levels 2-1 elements*. Every mapping can be adapted and modified by deleting/adding new elements (according to those accepted under the parent element) or modifying the elements properties.

The leaf meta-classes from figure 2 (dark elements) sketches some examples of future elements. They are on purpose not showing their attributes (for ease of reading). However each of them owns specific properties in accordance with the different in-progress formal specifications we are studying about the Moodle instructional design semantics, pedagogical activities and patterns, and activity structures.

The current abstract syntax proposition still has to be improved in order to allow the declaration of didactic objectives to the various *Level 4-to-1 elements*. Such objective will be mapped into Moodle *Objectives*, attached to the root *Course* and referenced by the direct or indirect corresponding *Level 1 elements*. Similarly, roles or groups have to be included in order to allow the division of labour in the learning scenario. Mappings to the Moodle concepts of *Group* and *Grouping* will be studied.

## 4 Illustration

### 4.1 A Learning Scenario Example

We on purpose propose to illustrate our proposal by formalising a very simple but representative learning scenario for the Moodle LMS. We propose at first a brief textual description, then the equivalent specification as a model conformed to the dedicated metamodel we proposed in section 3 (Figure 3 is a screendhot of the EMF-tree-based model editor, annotated to highlight the elements' level).

The learning scenario is composed of two learning sessions. The first one is a *lecture* session for which the teacher only want to propose learners with a

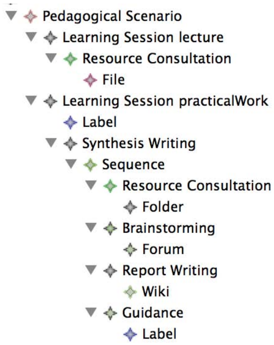



**Fig. 3.** Example of learning scenario composed of elements from the 4 levels

*Resource consultation* corresponding to his face-to-face course material. This pedagogical activity has the *quantity* property set to "one" and the location one set to "local". These properties will lead the dynamic mapping process to propose the *File* Moodle element.

The second learning session is a *practical work* that the teacher wants to realise in face-to-face within a computerized classroom. He would like to use the Moodle platform for supporting a pedagogical pattern "*Synthesis writing*". This pattern is automatically composed of a *sequence* activity structure embedding 4 sub-components. The first one is another *Resource consultation*. This time, the properties set to "many" (quantity) and "local" (location) by the teacher will lead the transformation process to add a *Folder* tool. The second sub-element is a *Brainstorming* pedagogical activity. Its *orientation* property set to "discussion" leads to propose a Forum tool. Similarly the third one is another pedagogical activity *Report writing* leading to a *Wiki* tool because of the *collaborative* property set to "true". Finally the fourth sub-component is a *Guidance* activity that aims at reminding the teacher to evaluate the synthesis in the wiki. Thanks to a *public* property set to "tutor" it leads the mapping process to set the corresponding *Label* to be invisible (*visible*="false") to students (it will be only visible to the teacher).

The teacher can change at any time the activities properties, leading to other mapping adaptations. He can also manually delete the mapping elements, rearrange their order, or add some other elements. Figure 3 shows a global overview of the learning scenario elements including all the automatic mappings according to the various properties and values (not depicted within the figure).

### 4.2 An Internal Mapping Example

According to our Model Driven approach, we can use model transformations to achieve such mappings. The transformations will be run on demand at design-time, to add mapped elements to the model and populate the sub-diagrams. Such transformations are complex (proportionally to the mapping complexity) and numerous, thus costly to write. We on purpose propose to use model weaving

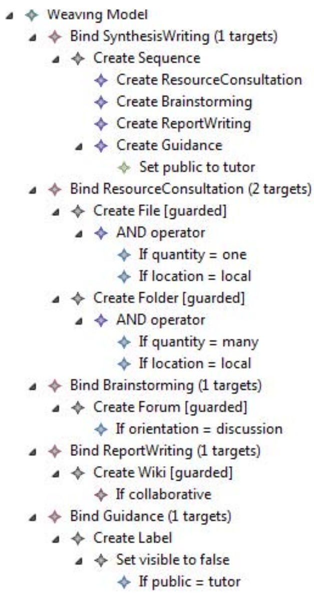



**Fig. 4.** Example weaving model specifying mappings from Figure 3

to capture the mapping semantics in dedicated weaving models and automatically generate models transformations. From a practical point a view, thanks to representative specifications from the teaching community, an engineer will model the mappings in a weaving model, using a tree based editor. He can then run a generic *High Order Transformation* (HOT) that will generate the concrete "mapping transformations". These final transformations can then be integrated within the graphical editor to be run at design-time.

The weaving models can be expressed using a weaving language, based on a generic weaving metamodel we designed. This weaving metamodel defines the "syntax" of the mapping/weaving model. Each mapping (or binding) has one *source* element and one or several *targets* (chosen from the extended instructional design metamodel). Targets can have conditions on whether they have to be instantiated or not, attributes can be set to specific values (also with conditions)... Figure 4 is a weaving model, displaying the mapping strategy from the example in section 4.1.

We used languages and tools from the Epsilon project to build a software framework fulfilling our model weaving requirements. Weaving models are edited through ModeLink, a three pane editor displaying the source and target metamodels in side panels (which are the same in our use case). The final "mapping" transformations are expressed using Epsilon Object Language (EOL), and are generated through a Model-to-text transformation using EGL language. This last transformation replaces the HOT traditionally used in model weaving environments.

## 5 Conclusions

This paper proposes a specific LMS-centered approach for raising the pedagogical expressiveness of its implicit learning design semantics. We discussed how the LMS low-level parameterizations could be abstracted in order to build higher-level building blocks. Based on the Moodle application, we present and illustrate our approach by formalising the abstract syntax of a Moodle-dedicated instructional design language following a specific 4-levels architecture. Such abstraction of LMS semantics may be a promising approach to develop a new generation of LMS-centered learning design languages, enabling teachers to develop pedagogically sound and technically executable learning designs.

The complete version of our metamodel proposition will drive the definition of the concrete syntax model (graphical notation), the palette and the mappings models in order to develop and tool the authoring-tool. Because of our former experiences about the EMF/GMF frameworks, we will also have to pay attention to the abstract syntax adjustments required in order to realise some specific visual representations.

We are also currently experimenting different frameworks about weaving and transforming models (more broadly about models composition). Indeed, the different default mappings during the design-time require a contextualised transformation model to perform. We are studying some weaving tools that will allow us to specify the mappings and automatically generate transformation rules (at design-time). First results have been illustrated within this article.

Also, in our approach the 4-levels extended metamodel will not allow to serialize future learning scenarios in conformance with the LMS format (source metamodel): a global transformation is required to restore this conformance. This transformation will be available as an export feature from our authoring-tool.

**Acknowledgments.** This article is part of the GraphiT project, a 42-months funded project of the French research agency. We thank the various individuals that participated to the interviews and surveys. We also thank the different project members involved in the identification and formalization task that help us in this publication.

## References

1. Abdallah, F., Toffolon, C., Warin, B.: Models transformation to implement a Project-Based Collaborative Learning (PBCL) Scenario: Moodle case study. In: 8th IEEE International Conference on Advanced Learning Technologies, pp. 639–643. IEEE Computer Society, Washington DC (2008)
2. Abedmouleh, A., Oubahssi, L., Laforcade, P., Choquet, C.: An analysis process for identifying and formalizing LMS instructional language. In: 7th International Conference on Software Paradigm Trends, pp. 218–223. ScitePress (2008)

3. Alario-Hoyos, C., Munoz-Cristobal, J.A., Prieto-Santos, L.P., Bote-Lorenzo, M.L., Asensio-Perez, J.I., Gomez-Sanchez, E., Vega-Gorgojo, G., Dimitriadis, Y.: GLUE! - GLUE!-PS: An approach to deploy non-trivial collaborative learning situations that require the integration of external tools in VLEs. In: 1st Moodle Research Conference, Heraklion, pp. 77–85 (2012)
4. Bergin, J., Eckstein, J., Manns, M.L., Sharp, H., Chandler, J., Marquardt, K., Wallingford, E., Sipos, M., Völter, M.: Pedagogical Patterns: Advice For Educators. Joseph Bergin Software Tools (2012)
5. Berggren, A., Burgos, D., Fontana, J.M., Hinkelman, D., Hung, V., Hursh, A., Tielemans, G.: Practical and Pedagogical Issues for Teacher Adoption of IMS Learning Design Standards in Moodle LMS. Journal of Interactive Media in Education, special issue: Advances in Learning Design (2005)
6. Bertelsen, O.W., Bodker, S.: Activity Theory. In: Carroll, J.M. (ed.) HCI Models,Theories and Frameworks: Toward a Multidisciplinary Science, pp. 291–324. Morgan Kaufmann, San Francisco (2003)
7. Burgos, D., Tattersall, C., Dougiamas, M., Vogten, H., Koper, R.: A First Step Mapping IMS Learning Design and Moodle. Journal of Universal Computer Science 13, 924–931 (2007)
8. Conole, G., Dyke, M., Oliver, M., Seale, J.: Mapping pedagogy and tools for effective learning design. Computers & Education 4, 17–33 (2004)
9. Dougiamas, M., Taylor, P.: Moodle: Using Learning Communities to Create an Open Source Course Management System. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications, pp. 171–178. Association for the Advancement of Computing in Education, Waynesville (2003)
10. Eclipse Modeling Project Official Website, http://www.eclipse.org/modeling/
11. Garrisson, D.R., Kanuka, H.: Blended learning: Uncovering its transformative potential in higher education. The Internet and Higher Education 7, 95–105 (2004)
12. Heathcote, E.A.: Learning design templates - a pedagogical just-in-time support tool. In: Minshull, G., Mole, J. (eds.) Designing for Learning, pp. 19–26. JISC Development Group (2006)
13. Katsamani, M., Retalis, S., Boloudakis, M.: Designing a Moodle course with the CADMOS learning design tool. Educational Media International 49, 317–331 (2012)
14. Loiseau, E., Laforcade, P.: Specification of learning management system-centered graphical instructional design languages - A DSM experimentation about the Moodle platform. In: 8th International Joint Conference on Software Technologies, pp. 504–511. Scitepress (2013)
15. Moodle Official Website, https://moodle.org
16. Ormrod, J.E.: Human Learning. Pearson College Division, Upper Saddle River (2011)

# Exploring the Potential of Speech Recognition to Support Problem Solving and Reflection

## Wizards Go to School in the Elementary Maths Classroom

Manolis Mavrikis[1], Beate Grawemeyer[2], Alice Hansen[1], and Sergio Gutierrez-Santos[2]

[1] London Knowledge Lab, Institute of Education, University of London, UK
{m.mavrikis,a.hansen}@ioe.ac.uk

[2] London Knowledge Lab, Birkbeck College, University of London, UK
{beate,sergut}@dcs.bbk.ac.uk

**Abstract.** The work described in this paper investigates the potential of Automatic Speech Recognition (ASR) to support young children's exploration and reflection as they are working with interactive learning environments. We describe a unique ecologically valid Wizard-of-Oz (WoZ) study in a classroom equipped with computers, two of which were set up to allow human facilitators (wizards) to listen to students thinking-aloud while having access to their interaction with the environment. The wizards provided support using a script and following an iterative methodology that limited on purpose their communication capacity in order to simulate the actual system. Our results indicate that the feedback received from the wizards did serve its function i.e. it helped modify students' behaviour in that they did think-aloud significantly more than in past interactions and rephrased their language to employ mathematical terminology. Additional results from student perception questionnaires show that overall students find the system suggestions helpful, not repetitive and understandable. Most also enjoy thinking aloud to the computer but, as expected, some find the feedback cognitively overloading, indicating that more work is needed on how to design the interaction tipping the balance towards facilitating post-task reflection.

**Keywords:** Speech Recognition, Wizard-of-Oz, Intelligent Support.

## 1 Introduction

The importance of language as both a psychological and cultural tool that mediates learning has been long identified; from as early as Vygotsky to modern linguists such as Pinker. From a Human Computer Interaction (HCI) perspective, speech recognition technology has the potential to enable more intuitive interaction with a system, particularly for young learners who reportedly talk aloud while engaged in problem solving (e.g. [1]). However, with the exception of limited research discussed in Section 2, the relationship between speech and

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 263–276, 2014.
© Springer International Publishing Switzerland 2014

learning has not been investigated in the Technology-Enhanced Learning (TEL) field, despite the advances that automatic speech recognition (ASR) is making.

The potential of such technology for learning relates to young children's capacity of inner-speech (that can become explicit think aloud) and reflection as a learning mechanism. For example, in mathematics, our area of interest, several researchers have emphasised that metacognitive instruction that uses self-directed speech improves students' mathematical reasoning [2]. Elementary school students can be aware of their inner-speech and verbalise their thoughts [1]. While the importance of explicit reflection is well understood [3], researchers have also suggested that reflecting is the answer to stimulating retention and that the skill of reflection must be taught at an early age [4].

In this paper we report studies performed in order to understand better the potential of speech-enhanced intelligent support in young children's interactive learning in the iTalk2Learn system, a platform that sequences tasks from intelligent tutoring and exploratory learning environments, while leveraging speech recognition technology to enhance the intelligent decision making, and speech production to provide feedback for students.

The main research question that concerns us: will students talk naturally to the computer and think-aloud, when seemingly unobserved? In the long term, we are interested in the potential of employing speech-recognition to enhance the system's ability to support student problem-solving, exploration and reflection.

We designed and conducted a Wizard-of-Oz (WOZ) study where wizards simulated realistic capabilities of speech recognition technology that shows promising results with respect to both encouraging students' verbal reflections and providing problem-solving support. To the best of our knowledge, the study was unique in its complexity and ecological validity in that it was conducted in a school classroom targeting in six sessions two students at a time while the rest of the classroom was working on the same system (with limited support) thus allowing the students to feel unobserved and think aloud and talk to the computer as they would normally do in the system proper.

In what follows, Section 2 presents a brief background with respect to the potential of voice interaction for learning. Section 3 outlines the methodology and the setting behind the WOZ studies. Section 4 presents descriptive results and an exploratory analysis of the effect of feedback on students' reactions. Section 5 discusses the results in more detail and Section 6 concludes the paper and draws implications for future research.

## 2 Background — Voice Interaction For Learning

From an HCI perspective speech production and recognition can provide potentially more intuitive interaction. In particular, spoken language input can enable students to communicate verbally with an educational application and thus interact without using human interface devices such as a mouse or keyboard.

Despite ASR for children being extremely difficult it is worth bearing in mind that related HCI-research suggests that ASR accuracy should not limit its usage and that the overall VUI design and the match of the application to its context should be able to compensate for possible flaws [5]. The approach taken in previous work (particularly the LISTEN [6] project) suggests that 100% accuracy can neither be expected nor relied upon. In light of this, it is preferential to err on the side of caution, thereby ensuring the least negative impact on learning.

With respect to learning in particular, the hypothesis that ASR can facilitate learning is based mostly on educational research that has shown benefits of verbalization for learning (e.g., [7,8,9]). The possible verbalization effect could be enhanced with ASR since cognitive theory of multimedia learning [10] predicts that a more natural and efficient form of communication will also have positive learning gains. The few existing research studies have found mixed results with respect to whether the input modality (speaking vs. typing) has a positive, negative or no effect on learning. In [11], for example, the authors investigated whether student typing or speaking leads to higher computer literacy with the use of AutoTutor. They reported mixed results that highlight individual differences among students and a relationship to personal preferences and motivation.

The importance of students' verbal communication in mathematics in particular becomes apparent if we consider that learning mathematics is often like learning a foreign language. Focusing, for example, on learning mathematical vocabulary, [8] encouraged students to talk to a partner about a mathematical text to share confusions and difficulties, make connections, put text into their own words and generate hypotheses. This way, students were able to make their tentative thinking public and continually revise their interpretations.

For further consideration is the research about self-explanation; an efficient learning strategy where students are prompted to verbalize their thoughts and explanations about the target domain to make knowledge personally meaningful. Previous research [12] found that the amount of self-explanation that students generated in a computer environment was suppressed by having learners type rather than speaking. Moreover, some students are natural self-explainers while others can be trained to self-explain [13]. Even when self-explanation is explicitly elicited, it can be beneficial [14] but requires going beyond asking students to talk aloud by using specific reflection prompts [13].

Self-explanation can be viewed as a tool to address students' own misunderstandings [14] and as a 'window' into students' thinking. While it may be early days for accurate speech recognition to be able to highlight specific errors and misconceptions, undertaking carefully-designed tasks can help identify systematic errors that students make. For example, [15] explores how naming and misnaming involves logic and rules that often aid or hinder students' mathematical learning and relate to misconceptions. A lack of mathematical terminology can also be noticed and prompts made to students to use appropriate language as they self-explain.

Finally, speech provides an additional cue for drawing inferences on students emotions and attitude towards the learning situation while they are solving tasks.

By paying attention to tone and pitch of speech in conjunction with other auditory signs like sighs, gasps etc., we can provide learners with even more individualized help, for instance, in the form of motivational prompts.

## 3 The Wizard-of-Oz study

### 3.1 Methodology

The studies reported on this paper are part of a process of a methodology referred to as Iterative Communication Capacity Tapering (ICCT) for designing the intelligent support for helping students in interactive educational applications [16]. In particular, they relate to both task-specific problem solving and affective support. ICCT is an extension of the well-known HCI wizard-of-oz methodology for the development of intelligent systems that recognises the complexity of educational contexts by advising a gradual reduction (tapering) of the communication between a human facilitator and the students followed by replacing the facilitator by a computer-based system. During the first phase, the facilitator gradually moves from a situation in which the interaction with the student is close, fast, and natural (i.e. face-to-face free interaction) towards a situation in which the interaction is mediated by computer technologies (e.g. voice-over-ip or similar for voice interaction, instant messaging or similar for textual interaction) and regularised by means of a script. On a second phase, the script is crystallized into a series of intelligent components that produce feedback in the same way that the human facilitator did. The gradual reduction of communication capacity and the iterative nature of the process maximise the probability of the computer-based support being as useful as the facilitator's help.In this paper, we are already starting the second phase, i.e. gradually replacing humans by a computer-based system. Experts ("wizards") are not physically near enough to the students to observe them directly, and therefore must observe them by indirect mediated means: the students' voice was heard by using microphones and headsets and their screen was observed by a mirror screen. The wizards did not have direct access to the students' screen (so e.g. could not point to anything on the screen to make a point), could not see the students' face (for facial cues), and could not communicate to students by using body language, only the facilities provided by the wizard-of-oz tools that resemble those of the final system.

### 3.2 Participants and Procedure

After returning informed consent forms signed by their parents 60 students 9 to 10-year old (Year-5) took part in a series of sessions with the iTalk2Learn platform configured for learning fractions through structured tasks from the intelligent tutoring system Whizz Maths and more open-ended tasks from the exploratory learning environment Fractions Lab. The sessions were designed to first familiarise all students with the environment and then allow them to undertake as many tasks as possible (in a study that has goals outside the scope of this paper). In parallel, we were running the WOZ study by asking two students in each session to work on different computers as described below. In total 12 students took part in the WOZ study but due to data errors we were able to analyse the interaction of only 10 students.

At the end of the session the students who participated in the WOZ study answered a questionnaire and took part in a focus group. Although methodologically students' opinion elicited through questionnaires can be problematic, they can still provide useful metrics that help us gauge students' perception of the intelligent support and influence our decisions in relation to the the overall approach we are taking [17]. We employed a 5-point Likert visual analogue scale with pictorial representations of smileys that children can relate to (see [17,18]) in order to respond to pertinent questions. We were particularly interested in *helpfulness*, *repetitiveness* and *comprehension*, metrics which we have previously identified useful [17]. In this research we added an extra construct that asked students whether they enjoyed thinking aloud.

### 3.3 Classroom Setup

The ecological validity of the study was achieved by following the setup depicted on Figure 1 and Figure 2. The classroom where the studies took place is the normal computer lab of the school in which more computers are on tables facing the walls on a Π-shape and a few are on a central table. This is the place where the WOZ study took place, while, for ecological validity, the rest of the class was working on the other computers. The students were told that the computers in the central isle were designed to test the next version of the system and were thus also responding to (rather than just recording as the rest of the computers) their speech. The central isle had two rows of computers, facing opposite directions, and isolated by a small separator for plugs etc. In the central isle the students worked on a console consisting on a keyboard, a mouse, and a screen. Usually, those components are connected to the computer behind the screen; for these studies, they were connected to a laptop on the wizards' side of the table. This allowed the wizard to observe what the students were doing. As the iTalk2Learn system is a web-based system, and all the students see is a web browser, the operating system and general look-and-feel of the experience was equivalent to the one that the rest of the students were using. When the wizards wanted to intervene, they used the iTalk2Learn WOZ tools to send messages to the student's machine. These messages were both shown on screen and read aloud by the system to students, who could hear them on their headset.

### 3.4 The Wizard'S Tools

In line with the ICCT methodology mentioned above, the wizards restricted their 'freedom' in addressing the students by employing a pre-determined script in which the expected interventions had been written. Figure 3 shows a high-level view of this script, the end-points of which require further decisions also agreed in advance in a protocol but not shown here for simplicity. In this study,

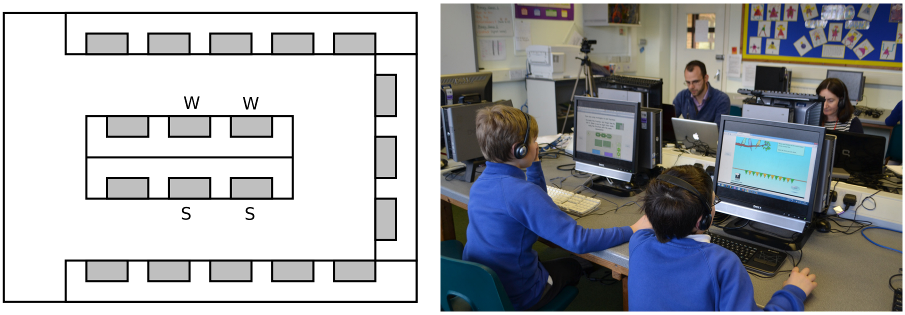



**Fig. 1.** The classroom. Most students are facing the walls on a Π-shape. The Wizard-of-Oz studies took place on the central isle while the rest of the students are working on a version of the system that only sequences tasks and provides minimal support.

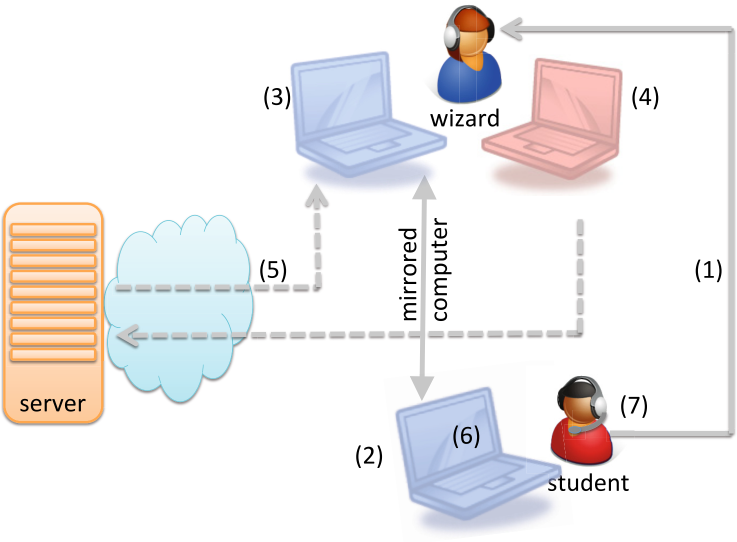



**Fig. 2.** Wizard-of-oz setup. Each student speaks on a headset (mic) that is connected to the wizard's headphones (1). The student interacts with a console (i.e. keyboard, mouse, screen) that is connected to a laptop on the wizard's side (2,3) so as the latter can witness their interaction. The wizard can send messages (4) by using specially designed wizard tools. These messages arrive to a server and subsequently to the mirrored laptop) (5) where they can be seen (6) and heard (7) by the student.

**Table 1.** Examples of feedback types

| Feedback type | Example |
|---|---|
| AFFECT | It may be hard, but keep trying.<br>If you find this easy, check you work and change the task |
| TALK ALOUD | Remember to talk aloud, what are you thinking?<br>What is the task asking you?' |
| TALK MATHS | Can you explain that again using the terms denominator, numerator? |
| PROBLEM SOLVING | You can't add fractions with different denominators. |
| REFLECTION | What did you learn from this task?<br>What do you notice about the two fractions? |

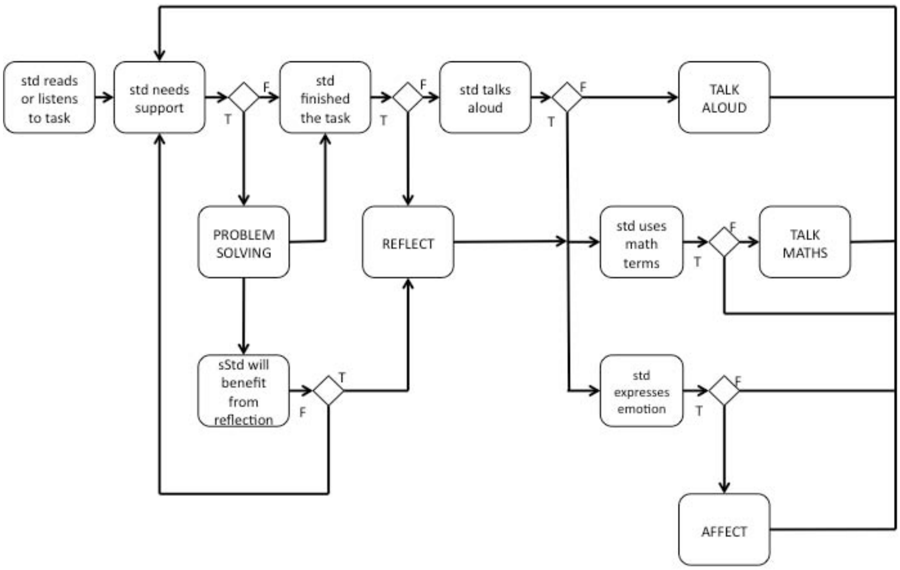



**Fig. 3.** Flowchart representing the wizard's script for support

we limited ourselves to written interventions that could be selected from an online document appropriate for being read aloud by the system. There were no other kinds of interventions (such as sounds, graphical symbols on screen etc.). The intervention had a set of associated conditions that would fire them thus resembling very closely the system under development.

## 4 Results

In total 170 messages were sent to 10 students. The raw video data was analysed by a researcher who categorised the feedback messages. The researcher noted

**Table 2.** Feedback types, including whether message was sent according to what student said

| | Was the feedback immediately related to what the student said? | | |
|---|---|---|---|
| Feedback type | NO | YES | Total |
| AFFECT | 15 | 16 | 31 |
| TALK ALOUD | 62 | 6 | 68 |
| TALK MATHS | 0 | 5 | 5 |
| PROBLEM SOLVING | 40 | 15 | 55 |
| REFLECTION | 9 | 2 | 11 |
| Total | 126 | 44 | 170 |

**Table 3.** Feedback types and whether student reacted immediate or not

| | Student reacted | | Response | | |
|---|---|---|---|---|---|
| Feedback type | NO | YES | Immediate | Delayed | Total |
| AFFECT | 8 | 23 | 19 | 4 | 31 |
| TALK ALOUD | 12 | 56 | 54 | 2 | 68 |
| TALK MATHS | 0 | 5 | 5 | 0 | 5 |
| PROBLEM SOLVING | 7 | 48 | 46 | 2 | 55 |
| REFLECTION | 0 | 11 | 10 | 1 | 11 |
| Total | 27 | 143 | 134 | 9 | 170 |

whether the feedback was directly related to what the student had said; and additionally whether the student reacted immediately, after a delay, or not at all. Another researcher went through the categories and any discrepancies were discussed and resolved before any analysis took place.

Table 1 shows the different types of messages sent to students. It can be seen that most frequent messages were reminders to talk aloud (68). This was followed by problem-solving feedback (55), and feedback according to students emotions (31). The least frequent messages relates to reflection (11) and using maths terminology (5).

By design, all of the reminders to use mathematical terminology were based on students' speech. In 16 of 31 cases, the feedback immediately related to what the student had said concerned emotions. It is not surprising that student reminders to talk aloud were not sent according to what the student had said, as they were provided when the student did not speak.

Students' reactions to the different feedback types can be seen in Table 2. Students reacted to all of the reminders to use mathematical terminology (100%) by repeating their sentence and making an effort to be more precise and use mathematical language such as 'fraction' (instead of 'that') and numerator, denominator (instead of 'this', 'top number', 'bottom number') as they often do. Additionally, students reacted to all feedback to reflect on the task (100%).This was followed by problem solving feedback (87%), and reminders to talk aloud (82%). The lowest number of reactions occurred after an affect boost (74%).

We investigated whether there is any correlation between feedback immediacy and response. There was no significant correlation between provision of feedback immediately after what the student had said, and students' reactions ($r=.18$, $p>.05$). However, the correlation indicates a positive trend on provision of feedback which is immediately following student verbalisation, followed by a reaction. Especially when we take the problem-solving support into account that does not rely on speech but on students' actions with the exploratory environment, there is a more clear correlation ($r=.16$, $p<.05$), as expected because of the several interventions on problem solving that do not necessarily require an immediate reaction from the student, but either to observe or think something that may not be directly observable.

We ran a one-way ANOVA on feedback types with respect to response types by categorising them as follows: affect-related, talk-aloud prompts, and learning-related (includes prompts to elaborate on the terminology, as well as specific problem solving feedback and reflection prompts). There was a significant effect of feedback type on immediacy of the response in that students were less likely to respond immediately after affect prompts $F(2,167)=4.05$, $p<.05$. We discuss this in light of the focus group and student questionnaire results in Section 5.

### 4.1 Student Questionnaire

Consistent with the findings in the literature and our previous studies, the students responded positively on the questionnaire and did not have any difficulties answering using the smiley visual analogue. Figure 4 depicts boxplots of the answers to the questionnaires that overall are quite satisfactory. One of the reasons for providing the questionnaire is that despite its limitations it can help us identify any negative perception that students may have, that tends not to be voiced in one-to-one or focus groups. As such, it was encouraging to see that all students find the feedback 'somewhat' or 'very' understandable. They also did not seem to find it repetitive, which is positive given the effort that we put in the design and the alternative texts that wizards have in their disposal even for the same meaning. We will discuss in more detail in the next Section why some students may not have found the support so helpful (even if we know that generally this is the lowest reported metric because what people, and young children in particular, perceive as helpful might not necessarily be aligned with what the pedagogical design suggests). The next section also discusses that the variety of answers with respect to whether students liked thinking aloud, even if positive overall, relates to individual differences and preferences.

## 5 Discussion

Designing intelligent educational systems is a complex endeavour that requires a holistic approach to both the system's behaviour and crucially the interaction with the student. The emphasis here is on interaction since (apart from the difficult task of providing problem-based support at the right time and level) it

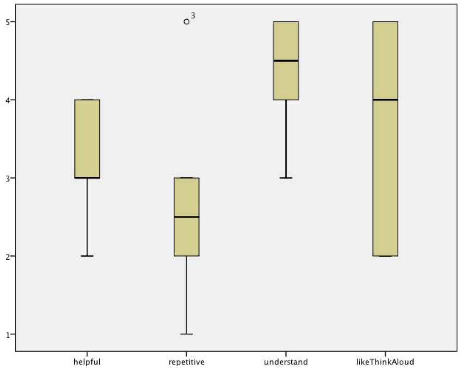



**Fig. 4.** Box plot of student answers

is important (especially during a WOZ study) to give student the appearance of an interaction that can help them believe both that it is adaptive to their interaction and that there is a benefit in, for example in our case, the effort they are putting in talking-aloud. The issues therefore are intertwined and relate both to the speech recognition capabilities (which here were simulated by wizards and thus had reduced error than an actual system) and feedback capabilities (which in our case was through the written and audio prompts). In what follows we categorise in themes the various issues that have emerged during our analysis including the focus group discussion with the students.

### 5.1 "Now You Are Talking" - The Effect of Ecological Validity

From a methodological perspective this study confirmed related literature and our anecdotal evidence that younger students are comfortable talking aloud when undertaking problem solving tasks. It also supported our conjecture that an ecologically valid setting would make a difference compared to earlier work in school with small individual or pair of students. In our previous visits students worked in a small room and at a single computer that was set up to receive WOZ messages. Anecdotal evidence showed that in this more artificial setup, students were less forthcoming thinking aloud despite prompts from the computer and/or the researcher. Prompts of the researcher to "Remember to think aloud" or to "Talk to the computer" were met with responses such as "I don't know what to say" or "I don't know how to talk to the computer". Some of the same students but even other ones who, according to the teacher's admission, are generally shy to talk, when in the authentic classroom setting talked more freely. This is in line with the self-explanation literature discussed in Section 2 that says that students are more likely to talk if they believe that this will help them solve the problem, rather than talking to an audience. There is a lot to attribute to the improved interaction with the computer (as compared to the previous, versions several things had improve to get the students a more realistic interaction).

However, while a proper statistical comparison is not appropriate in this case, we are convinced that the main difference was the ecologically valid setting i.e. that the students who were 'wizarded' felt less observed and therefore were more inclined to think-aloud.

### 5.2 "Sarcasm Doesn'T Become You" - Speech Production Matters

Students were positive about their experience at the computer overall. They liked receiving feedback as a direct result of their engagement with the tasks, particularly when it was reinforcing desired behaviour.

The students suggested improvements that could be made to the system. In earlier trials students in another school had heard feedback read out by a male voice. They did not like this voice, reporting it to be very brusque suggesting it sounded like they were being reprimanded each time. In order to improve the experience in this trial, an English female voice was used. The students felt that this voice had a sarcastic tone, particularly for feedback such as "well done". Furthermore, the students felt not all messages needed to be read out (particularly the feedback related to trying other problem-solving approaches), or that some could be stated but not shown as a pop-up because these tended to interfere with the flow of their work. This concurs with [13] findings when designing a multimedia environment that supports self-explanation by avoiding the duplication of messages across two different modalities (e.g. text and narration) that uses the same information processing channel. There was also a suggestion that students could choose when they wanted to read or hear feedback by having a button that they could click on when they needed help or wanted to hear what the computer was suggesting to them.

### 5.3 "Can the Headsets Look in My Head?" - Cognitive Load and AI

Most students, during the focus group, stated that they were at ease thinking aloud some even saying that it helped them concentrate. They were amused about the capabilities of artificial intelligence (AI) and provided (sometimes utopian) suggestions about what it can do. A couple however answered negatively in the questionnaire and in the comments raised a concern that they sometimes struggled to explain what they are doing. Although the latter is not a good justification for not thinking aloud (i.e. struggling to explain, hard as it may be, is definitely useful in clarifying one's thinking and contributes to learning as discussed in Section 2) we observed in some occasions that some of the students would get cognitively overloaded and lost in their own thinking aloud process in an effort to respond on the system's prompts. As the wizards were avoiding performing deep natural language processing they could not help the students. We observed similar situations in the recordings of the rest of the classroom that was encouraged to think aloud (in order to record their voices for training purposes of the ASR system) unaided. In a classroom of course this meant that a teacher or teacher assistant was able to provide more support. The focus group discussion revealed that a combination of factors might be affecting the students' perception on what they can or cannot do. Although all students undertook the same briefing, the fact that a computer can help them when talking aloud is not necessarily something that sinks in quite easily for all of them. This is exemplified by the student who muttered the title of this subsection to himself surprised by the feedback he received when after reflecting on a task he was prompted to repeat using mathematical terminology.

### 5.4 "Hmm—Let Me Rephrase That" - Pronouns 0 — Maths 1

We observed that most of the times the feedback served its function in that it modified the students' behaviour. The 143 out of 170 feedback prompts where student reacted is indicative of that. Albeit few, we are particularly interested in the prompts that requested rephrasing using mathematical terminology. We do not have the data to support further the conjecture but it seems to be a powerful enough prompt to help the students think more carefully what they are saying. One of the reasons for the low number of this type of feedback messages is that it was only needed once per student. The prompt seems to have been internalised by the students who (at least in one occasion) self-corrected herself replacing demonstrative pronouns with descriptive mathematical entities.

### 5.5 "Talking the Talk" - Individual Differences Matter

Individual differences matter in learning and interaction. Students' personality, affective characteristics and other factors play a role in their natural propensity to think aloud and/or to talk to an inanimate computer. Among the 12 students two were selected by their teacher on purpose as generally more silent in class and in the lower attainment group. Those received significantly lower feedback prompts related to their speech. We do not have enough data to support this statement but their teacher commented that even the few statements that the system (in the form of the wizard) elicited from them is an achievement. Furthermore, as with other students, we did observe them mumbling to themselves. So even if they were not talking aloud they seemingly engaged in inner-speech, which as mentioned in Section 2, has the potential to help reflection.

## 6 Conclusion

We presented an ecologically valid Wizard-of-Oz study designed to explore the potential of Automatic Speech Recognition (ASR) to support young students' exploration and reflection as they are working with interactive learning environments. The promising results indicate that, compared to our previous one-to-one settings, there is potential in expecting young students to think-aloud while interacting with educational technology especially if they see value by receiving support. Furthermore, even rudimentary ASR and decision-making mechanisms, as the ones presented here and simulated by our wizards, have the potential to support at the very basic level reflection on the learning task and on the use of the domain terminology. Additional information on affective state derived by cues either in the transcript or in the audio stream can also help in adapting problem solving support [19]. We conducted a cursory analysis on top-level categories that indicated a positive trend on students' more likely reaction on those feedback messages that are immediately following students verbalisation, evidence of a natural interaction. While students did not seem to react (at least in an observable manner) on affect-related prompts, the relationship between affect and feedback provision is very difficult to tease apart [20] and requires further research that was out of the scope of this paper.

We observed the positive reaction of the students to the (simulated) 'system' and collected their comments that feed to our next iteration. In particular, we identified a strong link between the quality of the speech production and the overall interaction with students' perception of the system and subsequent talk aloud. Our immediate steps in an effort to make the system more fluid is to improve the interaction by separating prompts that can only be read to students in contrast to those that have to be shown as well. Future work on the interaction will focus on how interruptive the messages are and carefully selecting the timings of the reflection prompts. We are leaning towards post-task reflection that is more natural and still has the capacity to contribute to student learning. In that sense we can utilise ASR in a non-intrusive manner and open the window to speech-enhanced intelligent support.

**Acknowledgments.** This research has received funding from the EU 7th Framework Programme (FP7/2007-2013 grant agreement No. 318051 – iTalk2Learn project http://www.italk2learn.eu). We would like to thank the school, students and teachers who took part in the study.

## References

1. Flavell, J.H., Green, F.L., Flavell, E.R., Grossman, J.B.: The Development of Children's Knowledge of Inner Speech. Child Development 68(1), 39–47 (1997)
2. Schoenfeld, A.H.: What's all the fuss about metacognition? In: Schoenfeld, A.H. (ed.) Cognitive Science and Mathematics Education, pp. 189–215. LEA (1987)
3. Kafai, Y.B., Harel, I.: Learning through consulting: When mathematical ideas, knowledge of programming and design, and playful discourse are intertwined. In: Harel, I., Papert, S. (eds.) Constructionism. Ablex Publishing Corporation (1991)
4. Freudenthal, H.: Major problems of mathematics education. Educational Studies in Mathematics, 133–150 (1981)
5. Canny, J.: The future of Human-Computer interaction. Queue 4(6), 24–32 (2006)
6. Mostow, J., Aist, G.: Evaluating tutors that listen: An overview of project listen. In: Feltovich, K.F.P. (ed.) MIT/AAAI Press, pp. 169–234. AAAI Press (2001)
7. Askeland, M.: Sound-based strategy training in multiplication. European Journal of Special Needs Education 27(2), 201–217 (2012)
8. Borasi, R., Siegel, M., Fonzi, J., Smith, C.: Using transactional reading strategies to support sense-making and discussion in mathematics classrooms: An exploratory study. Journal for Research in Mathematics Education 29, 275–305 (1998)

9. Mercer, N., Sams, C.: Tecahing children how to use language to solve maths problems. Language and Education 20(6), 507–528 (2007)
10. Mayer, R.E., Moreno, R.: Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist 38(1), 43–52 (2003)
11. D‘Mello, S.K., Dowell, N., Graesser, A.: Does it really matter whether student‘s contributions are spoken versus typed in an intelligent tutoring system with natural language? Journal of Experimental Psychology; Applied 2011 17(1), 1–17 (2011)
12. Hausmann, R.G.M., Chi, M.T.H.: Can a computer interface support self-explaining? Cognitive Technology 7(1), 4–14 (2002)
13. Roy, M., Chi, M.T.H.: The self-explanation principle in multimedia learning. In: Mayer, R.E. (ed.) Cambridge Handbook of Multimedia Learning, pp. 271–286. Cambridge University Press, New York (2005)
14. Chi, M.: Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In: Glaser, R. (ed.) Advances in Instructional Psychology, pp. 161–238. Lawrence Erbaum Associates, Mahwah (2000)
15. Furani, H.A.: Misconceiving or misnaming?: Some implications of toddlers' symbolizing for mathematics education. Philosophy of Mathematics Education Journal 17 (2003)
16. Mavrikis, M., Gutierrez-Santos, S.: Not all wizards are from Oz: Iterative design of intelligent learning environments by communication capacity tapering. Computers & Education 54(3), 641–651 (2010)
17. Mavrikis, M., Gutierrez-Santos, S., Geraniou, E., Noss, R.: Design requirements, student perception indicators and validation metrics for intelligent exploratory learning environments. Personal and Ubiquitous Computing 17(8), 1605–1620 (2013)
18. Read, J., Macfarlane, S.: Endurability, engagement and expectations: Measuring children's fun. In: Interaction Design and Children, vol. 2, pp. 1–23 (2002)
19. Janning, R., Schatten, C., Schmidt-Thieme, L.: Feature analysis for affect recognition supporting task sequencing in adaptive intelligent tutoring systems. In: EC-TEL 2014 (to appear 2014)
20. Porayska-Pomsta, K., Mavrikis, M., Pain, H.: Diagnosing and acting on student affect: the tutor's perspective. User Modeling and User-Adapted Interaction 18(1), 125–173 (2008)

# Social Network Sites as a Moderating Factor in E-Learning Platform Adoption: An Extension

Cayetano Medina[1] and Ramón Rufín[2]

[1] UNIR, Seville, Spain
cayetano.medina@unir.net
[2] UNED, Madrid, Spain
rrufin@cee.uned.es

**Abstract.** The use of Social Network Sites (SNS), and the improvements gained from new technologies through e-learning platforms are two of the main topics that are currently discussed in the area of higher education. Acquiring a better understanding of e-Learning platform adoption requires the addition of new antecedents to prevailing models and an analysis of their interrelations with other information and communication technologies. Web 2.0 and SNS are of particular significance in this connection. This paper attempts to determine the validity of the antecedents posited by Chan et al. (2010) in the field of e-Learning and the moderating role that SNS use may play in platform adoption. The influence of awareness, self-efficacy, convenience and assistance as antecedents of the components of the adoption model is determined; it is further found that SNS use has a moderating effect on the relationship between the intention to use the platform and its posited antecedents.

**Keywords:** e-learning, adoption, social network sites, UTAUT, awareness, self-efficacy, convenience, assistance, interaction effect, quasi-mediated relationships.

## 1 Introduction

The global acceleration of Information and Communication Technology (ICT) in education has been motivated by three developments in society: the exponential growth of knowledge which has meant that learners need to acquire the ability to process information; the second development is the increasing popularity of digital culture; the third is that learners can learn anytime and anywhere in an online learning context [1]. ICT dematerializes, delocates, and globalizes information, leading to a digitization process that changes the main format in which knowledge is conveyed, and this, in turn, changes the way people think and behave. By overcoming the limitations of time and space and customizing the learning environment to the specific individual needs and learning styles, ICT transforms the learning experience and increases the speed, flexibility, and efficiency of Knowledge transfer [1, 2].

The most widespread application of this concept is the use of e-learning platforms. In addition to many other benefits, web-based learning includes a change in methodology that promotes students' active participation, initiative, and critical thinking,

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 277–287, 2014.
© Springer International Publishing Switzerland 2014

with a fundamental shift in focus from teaching to learning [3]. The increasing popularity of e-Learning also has resulted in increased research interest in identifying the factors that underlie successful implementations of ICT interactive tools like any other information system its success depends on learners` acceptance of these technologies and the social interaction between them [4].

What factors drive undergraduate students` digital choice on whether to use technology for learning is important research because it helps to identify possible areas of support that educators need to provide in order to enhance students` technology use for learning [5]. In this regard, [6] present the role played by various antecedents of the key variables driving adoption in online environments: awareness, compatibility, self-efficacy, flexibility, aversion to personal interaction, trust, convenience and assistance. Their research relates to e-Government services, so it is doubtful whether their results can be generalized to e-Learning platforms. This paper accordingly sets out first to address the practicability of extending the antecedents of e-Government adoption considered to e-Learning platform adoption models [6].

The second element listed above is social interaction among e-Learning platform users. Since that is what will lead students to share knowledge, and since e-learning systems contribute to the development of virtual communities that provide an active, collaborative audience for creating content, it is important to consider if students' use of SNS -this use being considered as a personal attribute of students rather than an additional element of the e-learning platforms- will have an effect on their use of e-learning platforms. This will be the second objective of the present study.

## 2 Conceptual Framework

### 2.1 E-Learning Platform Adoption

One of the most widespread theories in technology acceptance analysis is the Unified Theory of Acceptance and Use of Technology (UTAUT), which posits the existence of four key elements for technology acceptance and use. It states that performance expectancy, effort expectancy, and social influence affect intention of use, while intention of use and facilitating conditions are what determine how the technology is actually used. According to previous theories, the main determinant of behavior is the intention to carry it out; in other words, actions are based on individual attitudes. Thus consumers are rational decision-makers who behave (actual use) according to the value attributed to the results of their behavior and to the expectations they have regarding that behavior in terms of achieving those results (intention of use).

Several studies have analyzed e-learning platform adoption based on the UTAUT model. While [7] concludes that the elements influencing behavior are limited to facilitating conditions and social influence, [8] hold the opposite view, stating that none of the listed components determine the intention to use e-learning. More recently, [9] has confirmed that performance expectancy and facilitating conditions influence behavior intention, whereas effort expectancy and social influence do not. In keeping with prior studies, this author also confirms that behavior intention determines e-learning platform acceptance [10, 11].

## 2.2 Identifying Antecedents of E-Learning Platform Adoption

It has been considered a range of antecedents of the UTAUT model based on the stages of the introduction of a new technology [6]. In this perspective, the preparatory stage emphasizes awareness; the segmentation stage is associated with compatibility and self-efficacy; the positioning stage relates to flexibility and aversion to personal interaction; finally, the execution stage is linked to trust, convenience and assistance.

Awareness that online services are available is essential to adoption, in so far as it bears upon both intention and use [12]. The more prevalent the advertising relating to online services, the greater the awareness of the people influencing our behavior [6].

Compatibility is the degree to which the new technology is consistent with past experience and modifies the user's behavior by facilitating the acquisition of information influencing the intention to adopt the new technology. In the educational context, in terms of the alignment between behavior and the values required of a potential adopter, compatibility is the critical determinant of e-Learning platform adoption, both directly and indirectly via perceived usefulness [4] [5] [9].

Self-efficacy is a judgment of a person's ability to use a technology and perform certain tasks, and relates to the degree to which the consumer perceives him/herself as an innovative person. It is a key antecedent of perceived cognitive effort (ease of use) associated with technology use [13]. Computer self-efficacy is a significant determinant of perceived usefulness and perceived ease of use, students with higher self-efficacy are likely to have more positive beliefs about the perceived usefulness and perceived ease of use of an e-Learning system [14].

Flexibility relates to the adaptability of technology components in terms of fitting the system to the consumer's needs [15]. This is a key element for technology adoption and use, because it leads to a better outcome by improving the response to new needs [6].

Aversion to personal interaction relates to the consumer's perception of the degree to which the technology enables him/her to access and use online services without having to interact with another person. People generally prefer self-service because it saves time and enhances personal control; these are key benefits for the perceived effectiveness of the service in relation to the perceived outcome [6] [16].

Consumers are more inclined to expect benefits from using a technology if they are confident it will work according to their expectations, so trust relates directly to performance expectancy [6] [17].

Convenience is the user's perception of the time and effort required to use the technology. People are more inclined to accept a technology if they think it will make their life simpler and easier. Although it has been found no significant effect of convenience on online service use [18], convenience is nonetheless a variable that merits consideration [6].

Assistance in the use of an online technology relates to the consumer's perception that he/she will be able readily and promptly to obtain help when he/she encounters any difficulty in using it. This is an essential component for technology in so far as it makes it easier to use, and is particularly significant when use is mandatory [6] [19].

Performance expectancy is determined by compatibility, flexibility, aversion to personal interaction and trust [6]. The antecedents of effort expectancy and facilitating conditions are the same – self-efficacy, convenience and assistance. Finally, the antecedent of social influence is awareness.

### 2.3 Use of Social Networks as a Moderating Factor in E-Learning Platform Acceptance

Furthermore, given that most of the key competences for lifelong learning are closely related to the skills generated by using 2.0 tools, it is essential to analyze the implications that are identified for some of their most widespread applications in terms of their effect on higher education.

SNS are defined as all the applications that increase group interaction, shared collaborative spaces, social connections, and information exchanges in a web-based environment. These sites brings people with common interests together and are particularly effective in promoting communication and collaboration, hence being quickly adopted in educational settings [20, 21, 22, 23].

Given that students complain about the lack of opportunities for establishing real communication, integrating SNS into educational practices has become a key element for achieving more robust learning opportunities [21]. The success of 2.0 learning systems is based on interactive and collaborative learning and in the development of online educational communities [22].

Some of the benefits offered by working SNS into the educational context are related to fostering socialization in a safe, practical environment; enabling inclusive education; reinforcing informal learning processes; supporting knowledge acquisition through subliminal learning processes; facilitating immersion in foreign language environments; reducing knowledge gaps; and contributing to boost and build a digital identity [20].

In higher education circles, the advisability of linking SNS to e-learning platforms is a subject of debate. Despite overall acceptance of their being essential at this point, results differ when the impact of SNS on the educational process is subjected to an empirical analysis. Among those who report a positive outcome, SNS reinforce collaborative learning, involve participants in critical reasoning, and boost communication and writing skills [20]. It has been suggested a positive correlation between SNS use and student involvement as predictors of academic success, as well as also helping students to adapt to the university's culture [23, 24]. As against this, other authors suggest that SNS can have a negative effect on school work when they are perceived differently in the educational and personal contexts [3] [25].

It is important to remember that although e-learning platform acceptance may depend on students' involvement and social interaction, without which it is doubtful whether the necessary exchange of knowledge could occur [26], SNS are ideal for providing an active, collaborative audience for content development by enhancing communication and cooperation between students [22] [23] [27]. Therefore, we have to ask ourselves whether students' use of SNS actually has a moderating effect on e-learning tool adoption, which would suggest the following hypothesis:

H1: SNS use has a moderating effect on the relationship between the antecedents of the intention to use an e-learning platform and that intention itself.

# 3 Methodology

## 3.1 Development of the Model

This study is based on the classic formulation of the UTAUT. The scales used for measuring the awareness, convenience, self-efficacy, flexibility, aversion to personal interaction, trust, assistance, performance expectancy, effort expectancy, social influence, and facilitating conditions constructs are drawn from [6], whereas intention of use and actual use are drawn from [34], all of them through three items. Lastly, frequency of social network use is measured according to one single item. Seven-value Likert scales are used throughout.

## 3.2 Fieldwork

Since the aim of this study was to measure the intention to use an e-learning platform, the questionnaire was personally handed out to students from the four undergraduate classes, to avoid a sample selection bias, and 155 valid questionnaires were obtained. As stated previously, SNS use is a personal attribute of students and not an additional element of the e-learning platforms, most of them actually using Facebook for that purpose.

The scales initially proposed for performing the statistical analysis were adaptations of scales that had been validated in several previous studies, and whose content we therefore could assume to be valid. The statistical behavior of the constructs included in the model was analyzed by developing the Structural Equations Model (SEM) with SmartPLS 2.0.M3 [28]. The purpose of PLS modeling is to predict dependent variables. This aim leads to an attempt to maximize explained variance (R2) of the dependent variables. Compared to covariance-based methods, PLS adapts better to prediction and theory development applications, although it can also be used for theory confirmation.

## 3.3 Measurement Model

Only three of the items used[1] displayed a factor loading less than 0.707 (SELF2: 0,659; API: 0,636; TRUS: 0,687), so the remaining items used exceeding that value and the required significance level[2] were retained.

---

[1] Abbreviations: AWAR (Awareness); COMP (Compatibility); SELF (Self-efficacy); FLEX (Flexibility); API (Aversion to Personal Interaction); TRUS (Trust); CONV (Convenience); ASSI (Assistance); PEXP (Performance expectancy); EEXP (Effort expectancy); SOCI (Social Influence); FACO (Facilitating Conditions); UINT (Intention of use); EUSE (Actual use).

[2] ($^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$ based on a one-tailed t-student (499) distribution, $t_{(0.05;\,499)} = 1.6479$; $t_{(0.01;\,499)} = 2.3338$; $t_{(0.001;\,499)} = 3.1066$.

### 3.4 Convergent and Discriminant Validity

Convergent validity is established by average variance extracted (AVE). It has been suggested that AVE values should exceed 0.5. This is the case here (Table 2), so the convergent validity of the related constructs can be accepted within the structural model. It has been recommended a value for R2 greater than 0.10 [29], in other case regards the values 0.67, 0.33 and 0.19 as substantial, moderate and weak, respectively [30]. Analysis of the values obtained shows that all variables exceed the weak level, except social influence. To establish the discriminant validity, the AVE value must be higher than the variance shared by the construct and the other represented constructs, and that condition was met in our study.

**Table 1.** Convergent validity

| | AVE | Composite Reliability | R2 | Cronbachs Alpha | Communality | Redundancy |
|---|---|---|---|---|---|---|
| API | 0.7091 | 0.8292 | - | 0.5995 | 0.7091 | - |
| ASSI | 0.6335 | 0.8383 | - | 0.7126 | 0.6335 | - |
| SELF | 0.7501 | 0.8572 | - | 0.6681 | 0.7501 | - |
| COMP | 0.8168 | 0.9304 | - | 0.8879 | 0.8168 | - |
| FACO | 0.7213 | 0.8857 | 0.3587 | 0.8065 | 0.7213 | 0.1849 |
| TRUS | 0.7125 | 0.8320 | - | 0.5981 | 0.7125 | - |
| AWAR | 0.7233 | 0.8868 | - | 0.8083 | 0.7233 | - |
| CONV | 0.6969 | 0.8731 | - | 0.7814 | 0.6969 | - |
| EEXP | 0.7150 | 0.8825 | 0.4510 | 0.8015 | 0.7150 | 0.2445 |
| PEXP | 0.7405 | 0.8954 | 0.4494 | 0.8247 | 0.7405 | 0.2096 |
| FLEX | 0.8546 | 0.9463 | - | 0.9149 | 0.8546 | - |
| SOCI | 0.8355 | 0.9384 | 0.1324 | 0.9013 | 0.8355 | 0.1106 |
| UINT | 0.8479 | 0.9436 | 0.2682 | 0.9102 | 0.8479 | 0.1997 |
| EUSE | 0.7350 | 0.8926 | 0.3133 | 0.8202 | 0.7350 | 0.1551 |

### 3.5 Structural Model

Analysis of the structural model establishes that compatibility (path: 0.2862, tvalue: 2.8362), aversion to personal interaction (path: 0.3884, tvalue: 5.1512) and trust (path: 0.2066, tvalue: 2.1437) exert a significant influence on performance expectancy, while flexibility does not (path: -0.0164, tvalue: 0.1897). The significant antecedents of effort expectancy are self-efficacy (path: 0.1566, tvalue: 2.1135), convenience (path: 0.2507, tvalue: 2.5519) and assistance (path: 0.4372, tvalue: 5.7817). The antecedent of social influence is awareness (path: 0.3639, tvalue: 4.9809). Self-efficacy (path: 0.2346, tvalue: 2.8108), convenience (path: 0.1687, tvalue: 1.7533) and assistance (path: 0.3705, tvalue: 4.1918) are the antecedents of facilitating conditions.

The significant antecedents of intention of use are effort expectancy (path: 0.3853, tvalue: 2.5654) and social influence (path: 0.1444, tvalue: 2.2496), but not performance expectancy (path: 0.0699, tvalue: 0.4450). Finally, intention of use (path: 0.3164, tvalue: 3.1287) and facilitating conditions (path: 0.3130, tvalue: 3.8813) are both antecedents of actual use of the e-Learning platform.

### 3.6 Quasi-Mediated Relationships

**Table 2.** Mediated relationships

| | |
|---|---|
| AIP⟶PEXP⟶UINT⟶EUSE | (-0.0366; 0.0459) |
| ASSI⟶ FACO⟶ EUSE | (0.0435; 0.2069) |
| ASSI ⟶ EEXP⟶ UINT ⟶ EUSE | (0.0056; 0.1149) |
| SELF⟶ FACO ⟶ EUSE | (0.0167; 01657) |
| SELF ⟶ EEXP ⟶ UINT ⟶ EUSE | (-0.0009; 0.0493) |
| COMP ⟶ PEXP ⟶ UINT⟶ EUSE | (-0.0218; 0.0518) |
| TRUS⟶ PEXP ⟶ UINT ⟶ EUSE | (-0.0229; 0.0208) |
| AWAR⟶ SOCI⟶ UINT ⟶ EUSE | (0.0012;0.0496) |
| CONV⟶ FACO ⟶ EUSE | (-0.0048; 0.1481) |
| CONV ⟶ EEXP ⟶ UINT ⟶ EUSE | (0.0008; 0.0897) |
| FLEX⟶ PEXP ⟶ UINT ⟶ EUSE | (-0.0087; 0.0130) |

Given the importance of mediation effects for this research model, they must be explored by systematic analysis [31]. Based on repeated samples, a confidence interval of 95% was generated for each mediator. If the interval does not contain 0, then the indirect effect of that mediator is significantly non-zero [31].

Our results (Table 2) enable us to establish that, out of the antecedents originally considered for the UTAUT model, awareness, self-efficacy, convenience and assistance exhibit a significant effect on the use of e-Learning platforms.

### 3.7 Interaction Effect

**Table 3.** Moderated relationships

| | |
|---|---|
| AIP ⟶PEXP | 0.0048 (0.0132) |
| ASSI ⟶FACO | -0.3606 (0.7707) |
| ASSI ⟶EEXP | -0.2939 (0.9451) |
| SELF⟶ FACO | -0.2157 (0.5479) |
| SELF ⟶ EEXP | 0.0075 (0.0222) |
| COMP⟶ PEXP | 0.4549 (1.1526) |
| FACO ⟶EUSE | 0.1387 (0.3197) |
| TRUS⟶ PEXP | 0.3236 (0.7307) |
| AWAR⟶SOCI | -0.0516 (0.1459) |
| CONV⟶ FACO | -0.9597 (2.3980)** |
| CONV ⟶ EEXP | -0.5005 (1.3242) |
| EEXP ⟶UINT | -0.9959 (2.3575)** |
| PEXP ⟶ UINT | -1.2769 (3.0275)** |
| FLEX⟶PEXP | -0.0358 (0.0979) |
| SOCI ⟶ UINT | -0.8526 (2.4827)** |
| UINT ⟶ EUSE | 0.1475 (0.3716) |

To confirm whether students' use of SNS had an effect on e-learning platform adoption, we proceeded to analyze the interaction effect. It is the most straightforward method for analyzing moderating effects, and its use is recommended whenever possible [32].

As to the moderating effect of the use of SNS on the relations driving adoption of e -Learning platforms, we conclude that this factor moderates the relationship between convenience and facilitating conditions and the relationships holding between performance expectancy, effort expectancy, social influence and intention of use.

# 4 Analysis of the Results

The first hypothesis to confirm was the one that analyzed the potential moderating effect on the relationship between intention of use and its antecedents. As we can see from the interaction effect, SNS use among students has a moderating effect on the relationship that performance expectancy, effort expectancy, and social influence have on intention of use, and on the relationship between convenience and facilitating conditions, leading us to accept H1. The resulting value shows that the greater the use of social network sites, the lesser the influence of these variables, although in the structural model performance expectancy does not appear to have a significant effect on the intention to use an e-learning platform.

# 5 Conclusions

## 5.1 Theoretical Implications

The present study establishes that students' intention to use an e-learning platform is determined by effort expectancy and social influence, but not by performance expectancy. In keeping with previous studies, we found that social influence had a strong effect on intention of use because, even though that use was not mandatory. Social motivations thus influence the development of positive attitudes toward the use of e-Learning platforms [4]. In keeping with [33], effort expectancy is a critical determinant for the intention to adopt e-learning. Contrary to what some authors have argued [8], performance expectancy not only does not have a determining influence on behavior intention, but does not even have a significant effect on it.

In order to further pursue the interpretation of Lin`s approach [35], according to which effort expectancy plays a more prominent role among inexperienced users, the experience that users of these platforms have ought to have been explicitly quantified. In other words, although according to the students' own responses in some cases they are indeed experienced, it is difficult to determine the source of this experience. Although some respondents may have acquired it at other universities, it seems like a reasonable assumption that a large percentage of these students would lack that experience. Finally, in keeping with prior studies, we found that intention of use and facilitating conditions determine university students' actual use of e-learning platforms [7] [10] [11].

A further point of interest is to identify the antecedents of the main components of the adoption model. Effort expectancy is determined by students' self-efficacy, by the convenience of platform use, and by the assistance available. Social influence is determined by awareness, while facilitating conditions are determined by self-efficacy, convenience and assistance. Finally, performance expectancy has its significant antecedents in compatibility, aversion to personal interaction and trust, whereas flexibility does not exert a significant influence.

If we focus on the role of these antecedents in determining use, we find that use is determined by assistance, self-efficacy, awareness and convenience. The determining role of convenience places our conclusions in opposition to previous works [18], and in alignment with other [6]. A particularly substantial role is played by assistance, which affects the actual use of the platform both via facilitating conditions and via effort expectancy [13]. Awareness of the e-Learning platform bears both upon intention of use and actual use [12]. In addition, the role of convenience is found to be decisive even for non-mandatory services.

The success of e-learning depends on students' participation, involvement, and social interaction, which could be reinforced by SNS use, this use being considered as a personal attribute of students and not an additional element of the e-learning platforms [26]. SNS use is thus independent from the educational institutions promoting the e-learning platforms. The present study concludes that SNS use moderates the relationship between the antecedents for behavior intention and that very intention, but not the effect of behavior intention and facilitating conditions on actual resulting use of the platform.

SNS use leads performance expectancy, effort expectancy, and social influence to have a weaker effect on behavior intention. Greater SNS use may imply that the student is familiar with the technology, and therefore, as posited in previous studies [33] [35] the greater the user's experience is, the weaker the impact of effort expectancy will be. Getting students involved in active interaction based on SNS can be one of the ways to change users' perceptions so as to have the lesser impact on effort expectancy that we mentioned earlier.

What is the real impact of SNS use on e-Learning platform adoption? Since all relationships among the antecedents of the intention to use an e-Learning platform are moderated, the only unmoderated relationships are those linking facilitating conditions to e-Learning platform use. But since the use of SNS moderates the relationship between convenience and facilitating conditions, the only unmoderated relationship is that holding between facilitating conditions on one hand and self-efficacy and assistance on the other. Both these variables, via mediated relations, bear upon e-Learning platform use, so we can conclude that some elements are unaffected by the impact of SNS.

The socialization and enculturation of young people in virtual environments, and the development of certain technological literacy-related skills pertaining to SNS use with a clearly critical, collaborative, and creative character lead to differing behavior of the relationship between social influence and the intention to use an e-learning platform, which results in social influence having a lesser impact on intention of use as its use increases. It is important to bear in mind that, according to our findings, SNS use has a positive effect on of the intention to use e-learning platforms.

The relationships that remain unaffected are those that relate the antecedents for intention of use to actual use. In other words, the significant effect of intended use and facilitating conditions on actual use of an e-learning platform is not dependant on students' SNS use. Furthermore, out of the identified antecedents of the main elements of the model, social network use only moderates the relationship between convenience and facilitating conditions.

## References


1. Uzunboylu, H., Bicen, H., Cavus, N.: The efficient virtual learning environment: A case study of web 2.0 tools and Windows live spaces. Computers & Education 56, 720–726 (2011)
2. Roca, J.C., Chiu, C.-M., Martínez, F.J.: Understanding e-Learning continuance intention: An extension of the technology acceptance model. International Journal of Human-Computer Studies 64, 683–696 (2006)
3. Cole, M.: Using Wiki technology to support student engagement: Lessons from the trenchs. Computers & Education 52(1), 141–146 (2009)
4. Hernández, B., Montaner, T., Sese, F.J., Urquizu, P.: The role of social motivations in e-learning. How do they affect usage and success of ICT interactive tools? Computers in Human Behavior 27, 2224–2232 (2011)
5. Lai, C., Wang, Q., Lei, J.: What factors predict undergraduate students' use of technology for learning? A case from Hong Kong. Computers & Education 59, 569–579 (2012)
6. Chan, F.K.Y., Thong, J.Y.L., Venkatesh, V., Brown, S.A., Hu, P.J.-A., Tam, K.T.: Modeling Citizen Satisfaction with Mandatory Adoption of an E-Government Technology. Journal of the Association for Information Systems 11(10), 519–549 (2010)
7. Teo, T.: Modelling technology acceptance in education: a study of pre-service teachers. Computers & Education 52(1), 302–312 (2009)
8. McGill, T.J., Klobas, J.E.: A task-technology fit view of learning management system impact. Computers & Education 52, 496–508 (2009)
9. Chen, J.-L.: The effects of education compatibility and technological expectancy on e-learning acceptance. Computers & Education 57, 1501–1511 (2011)
10. Liao, H.L., Lu, H.P.: The role of experience and innovation characteristics in the adoption and continued use of e-learning websites. Computer & Education 51(4), 1405–1416 (2008)
11. Dulle, T.W., Minishi-Majanja, M.K.: The suitability of the Unified Theory of Acceptance and Use of Technology (UTAUT) model in open Access adoption studies. Information Development 27(1), 32–45 (2011)
12. Toral, S.L., Barrero, F., Martínez-Torres, M.R.: Analysis of utility and use of web-based tool for digital signal processing teaching by means of a technological acceptance model. Computers & Education 49(4), 957–975 (2007)
13. Argawall, R., Sambamurthy, V., Stair, R.M.: Research Report: The evolving Relationship Between General and Specific Computer Self-Efficacy –An Empirical Assessment. Information Systems Research 11(4), 418–430 (2000)
14. Chow, M., Herold, D.K., Choo, T.-M., Chan, K.: Extending the technology acceptance model to explore the intention to use Second Life for enhancing healthcare education. Computers & Education 59(4), 1136–1144 (2012)
15. Mueller, D., Strohmeier, S.: Design characteristics of virtual Learning environments: State of research. Computers & Education 57, 2505–2516 (2011)
16. Meuter, M.L., Ostrom, A.L., Roundtree, R.I., Bitner, M.J.: Self-service technologies: Understanding customer satisfaction with technology-based service encounters. Journal of Marketing 64(3), 50–64 (2000)
17. Sok, Y., Chan, F.: Internet Banking Adoption in Kuala Lumpur: An Application of UTAUT Model. International Journal of Business and Management 6(4), 161–167 (2011)
18. Islam, M.A., Akkaruzzannon, M., Ramayah, T.: The Adoption of Mobile Commerce Service among Employed Mobile Phone Users in Bangladesh: Self-efficacy as a moderator. International Business Research 4(2), 80–89 (2011)

19. Hsieh, J.J.P.A., Rai, A., Keil, M.: Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. MIS Quarterly 32(1), 97–126 (2008)
20. Ajjan, H., Hartshorne, R.: Investigating faculty decisions to adopt Web 2.0 technologies: Theory and empirical test. The Internet and Higher Education 11(2), 71–80 (2008)
21. Mazman, S.G., Usluel, Y.K.: Modeling educational usage of Facebook. Computers & Education 55, 444–453 (2010)
22. Wang, H.C., Chiu, Y.T.: Assessing e-learning 2.0 system success. Computers & Education 57, 1790–1800 (2011)
23. Zhan, Z., Xu, F., Ye, H.: Effects of an online Learning community on active and reflective learners‘ learning performance and attitudes in a face-to-face undergraduate course. Computers & Education 56, 961–968 (2011)
24. Kuh, G.D.: What student affairs professionals need to know about student engagement. Journal of College Student Development 50(6), 683–706 (2009)
25. Junco, R., Cotton, S.R.: Perceived academic effects of instant messaging use. Computers & Education 56, 370–378 (2011)
26. Ma, W.W.K., Yuen, A.H.K.: Understanding online knowledge sharing: An interpersonal relationship perspective. Computers & Education 56, 210–219 (2011)
27. Harrinson, T.M., Barthel, B.: Wielding new media in Web 2.0 explaining the history of engagement with the collaborative construction of media products. New Media & Society 11(1-2), 155–178 (2009)
28. Ringle, C.M., Wende, S., Will, A.: SmartPLS 2.0 (Beta). University of Hamburg, Hamburg (2005), `http://www.smartpls.de` (retrieved 2008)
29. Falk, R.F., Miller, N.B.: A Primer for soft modelling. University of Akron Press, Akron Ohio (1992)
30. Chin, W.W.: The partial least squares approach to structural equation modeling. In: Marcoulides, G.A. (ed.) Modern Methods for Business Research, pp. 295–336. Lawrence Erlbaum, Mahwah (1998)
31. Sun, H.S.: Trust and Continued Use of Online Marketplaces. Journal of the Association for Information Systems 11(4), 182–211 (2010)
32. Rigdon, E.E., Schumacker, R.E., Wothke, W.: A comparative review of interaction and nonlinear modeling. In: Schumacker, R.E., Marcoulides, G.A. (eds.) Interaction and Nonlinear Effects in Structural Equation Modeling, pp. 1–16. Lawrence Erlbaum, Mahwah (1998)
33. Wang, Y.-S., Wu, M.-C., Wang, H.-Y.: Investigating the determinants and age and gender differences in the acceptance of mobile learning. British Journal of Educational Technology 40(1), 92–118 (2009)
34. Wang, Y.-S., Shih, Y.-W.: Why do people use information kiosks? A validation of the United Theory of Acceptance and Use of Technology. Government Information Quarterly 26, 158–165 (2009)
35. Lin, K.-M.: E-Learning continuance intention: Moderating effects of user e-learning experience. Computers & Education 56, 515–526 (2011)

# Adapting Tutoring Feedback Strategies to Motivation*

Susanne Narciss[1], Sergey Sosnovsky[2], and Eric Andres[2]

[1] University Passau, Lehrstuhl fuer Psychologie,
Dr. Hans-Kapfinger Str. 14b, 94030 Passau, Germany
Susanne.Narciss@uni-passau.de
[2] German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
{sergey.sosnovsky,eric.andres}@dfki.de

**Abstract.** This paper investigates tutoring feedback strategies adaptive to student motivation. Several static feedback adaptation strategies have been designed based on the interactive tutoring feedback (ITF) model, implemented within the Adaptive Educational System (AES) ActiveMath and evaluated in a study with 6th and 7th graders. Student motivation profile (high vs. low perceived competence and intrinsic motivation) has been used to assign them to either conceptual or procedural feedback condition. The data analysis shows that, for low motivated students, the benefits of adapting feedback to motivational profiles are visible in both, performance during treatment and knowledge gain from pre- to post-test. For highly motivated students, no significant effects have been registered. These findings shed light on the role of motivation in tutoring feedback processing and have important methodological implications for designing feedback strategies in AESs.

## 1 Introduction

Effective feedback is one of the most powerful factors influencing learning in various instructional contexts including technology-enhanced learning (TEL) environments (e.g., [1, 2]). The term "*feedback*" refers to all post-response information helping students regulate further learning [3, 4]. This notion of feedback can be traced back to early cybernetic views of feedback and implies that the core aim of feedback is to eliminate the gap between the current and the desired states of learning [5, 6].

The interactive potential of modern TEL environments makes it possible to provide students not only with basic feedback types merely communicating the outcome of a student response, but also with tutoring feedback guided by strategies. *Tutoring feedback strategies* combine formative elaborated feedback (such as hints, explanations, attribute-isolation examples) with instructional design methods (e.g., mastery learning or scaffolding) to support students in acquiring target competencies, detecting errors, overcoming learning obstacles, etc. In doing so, *tutoring feedback strategies* offer

* This work has been partly supported by the German National Science Foundation (DFG – Deutsche Forschungsgemeinschaft; Grants ME 1136/8-1 and NA 738/10-1).

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 288–301, 2014.
© Springer International Publishing Switzerland 2014

"*strategically*" useful information for task completion without immediately providing the correct solution [3, 12].

There is a fast-growing body of empirical research on feedback adaptation. Corbalan, Paas and Cuypers investigated the effect of different feedback levels (after each step vs. on the final step) in basic linear algebra tasks on perception, learning and motivation [7]. They found that students in general tend to prefer fine-grained feedback on each step and suggested the need for future studies on adapting the feedback level to individual student expertise. Mitrovic, Ohlsson and Barrow evaluated the impact of providing positive feedback in a constraint-based tutor that previously delivered only post-error feedback [8]. Their findings indicate that the inclusion of positive feedback can lead to faster learning. The research of Dennis, Masthoff, and Mellish shows that students themselves consider that adapting feedback may be not needed for successful students, but can be very important when a student is underachieving [9]. Conati and Manske explored the effectiveness of adaptive feedback in a serious learning game. Their results suggest that when feedback is too frequent and more informative than the students need, its effectiveness decreases [10]. Goldin, Koedinger and Aleven data-mined several models measuring individual feedback value and found that, for different students, the value of feedback messages is different; this allowed them to hypothesise about the existence of a feedback processing skill that defines the individual effectiveness of feedback depending on the level of provided details [11].

Most of the existing research, however, focuses on adapting feedback to student cognitive and meta-cognitive characteristics. The purpose of this work is to investigate how tutoring feedback strategies can be adapted to student motivation. In the reminder of this paper, we present an experimental study, which explores if and under what conditions students with different motivational profiles benefit from different tutoring feedback strategies.

## 2 The Context of This Work

### 2.1 Interactive Tutoring Feedback Model

The ITF-model considers feedback as a multidimensional instructional activity helping students to regulate their learning process and assisting them in acquiring knowledge and competencies needed to perform learning tasks. Such conceptualization of feedback facilitates analysis of possible factors and effects of (tutoring) feedback strategies [2, 3, 12]. Rooted in regulatory paradigms from the systems theory, the ITF-model views feedback as one of the several basic components of a generic feedback loop and suggests to differentiate between the two interacting feedback loops: the student feedback loop and the loop of the external feedback source. The controlled process for the two interacting feedback loops consists of the acquisition of competencies necessary to master the requirements of learning tasks. Building on models of self-regulated learning (e.g., [13]) and on current multidimensional notions of competencies (e.g., [14]), the ITF-framework suggests to determine the controlled variables for a learning process with regard to the cognitive, motivational, and meta-cognitive variables that are relevant for acquiring the desired competencies.

To regulate and control the acquisition of competencies, both feedback loops need to determine a standard or reference level for each of the controlled variables. Furthermore, it is assumed that the actual state of competencies is (continuously) assessed by both the student (internal assessment) and the external feedback source (external assessment). These assessments result in the internal and external feedback on the actual state of the student's competencies. In both feedback loops, the feedback on the current state of competencies has to be compared to the respective desired level of competencies in order to evaluate to what extent the desired level of competencies has been achieved. Based on this comparison, the external controller (e.g., an AES or a teacher) generates an external feedback message. If a gap between the desired and the current level of competencies has been detected, this external feedback message may provide evaluative information pertaining to this gap, as well as tutoring information (i.e., a suggestion of an action that would helps eliminating the gap).

The student has to process the external feedback along with the internal feedback in order to generate ideas or suggestions regarding one or several control actions that may help to proceed in the direction of the desired level of competency. Finally, the student has to implement the selected control action.

Given the complexity of this interaction between the feedback loops, the ITF-model suggests that the effects of tutoring feedback strategies depend not only on the properties of the feedback strategy, but also on various conditions and characteristics of the feedback receiver (i.e., a student), the parameters of the feedback source, and the instructional context [12]. An essential precondition for external feedback being efficient is the mindful processing of this feedback by the student.

### 2.2 AtuF Project

This work was a part of the project "Adaptive tutoring Feedback" (AtuF) that has the overall goal to systematically develop and evaluate adaptation strategies that allow tailoring tutoring feedback to students' needs. In the first project phase, the methodological, conceptual and technological framework for investigating tutoring feedback adaptation has been developed. Additionally, potential factors that should be taken into account when designing feedback adaptation strategies were investigated with an experimental study [15]. This paper presents the results of the 2nd project phase, which focused on the design and evaluation of static feedback adaptation strategies adjusting the content and sequence of feedback messages to student motivational profiles.

## 3 Experiment Description

### 3.1 The Domain of Fraction Arithmetic

The target domain of this research is the subset of basic fraction arithmetic taught in 6th and 7th grades of the German school program. We have employed the two-dimensional framework of fraction competencies [16] to represent target knowledge components and cognitive operations in the domain, as well as the typical errors students make while solving fraction exercises. All assessment and training exercises were modelled in terms of the elements of this framework. On the top level, exercises were grouped into the following five large categories based on the requirements addressed by the encoded typical error:

- Represent: Representation of fractions in various formats (e.g., candy bar, pie)
- Transform: Expansion and cancelling of fractions
- Add (common): Addition of fractions with common denominators
- Add (uncommon): Addition of fractions with uncommon denominators
- Order: Ordering and comparison of fractions

### 3.2 The Exercise Environment of ActiveMath

ActiveMath is a Web-based intelligent AES for mathematical domains [17]. It allows students to work on learning material at their own pace with various degrees of system support, ranging from intelligent problem solving support to personalized courses and adaptive navigation [18].

Interactive exercises in ActiveMath can have multiple steps and combine various types of tasks from single choice to formula input. On the representational level, ActiveMath exercises are finite state machines (FSM) consisting of nodes corresponding to (sub-)tasks and system's feedback, and transitions encoding correct and typical incorrect student responses [19]. Representations of an exercise can be automatically enriched by a tutoring strategy that defines a general pattern for reacting to the actions and requests of the student solving the exercise [20]. Technically speaking, a tutoring strategy is a transformation of a given exercise automaton into a new exercise automaton, which expresses the needed tutoring behaviour with respect to the instructional goals of the strategy.

Based on this framework, the AtuF exercise strategies can express different tutoring behaviour by varying the sequence and the types of feedback messages presented in a concrete instance of the exercise depending on various information stored in the student model. For example, the same exercise can offer a conceptual hint to a student with a high level of motivation and a procedural hint to a low motivated student.

### 3.3 Assessment Tests

In order to test target student competencies before and after training, two parallel sets of fraction assessment items were created. The items were designed by systematically combining conceptual and/or procedural knowledge components with one of the cognitive operations. The sets covered all of the suggested processes and knowledge types and their combinations with items of various response formats (e.g., multiple-choice, short-answer, constructed-response, and graphical input) [21].

To examine the psychometric properties of these tests, a dedicated pen-and-paper experiment was conducted with 293 6th and 7th graders. The analysis showed that the item difficulties ranged from .03 to .88. The difficulties were normally distributed for each set with a mean of .45 (pre-test) and .41 (post-test). For each set of tasks the mean discriminatory index was .31 (ranging from -.25 to .61). Reliability-analysis showed a Cronbach's alpha of .73 (pre-test) and .80 (post-test), containing only results of students who worked on every task.

Finally, we have selected two isomorphic subsets of these exercises for implementing in ActiveMath. The resulting pre- and post-tests consisted of 29 items each. While taking the tests within ActiveMath, students worked with series of exercises grouped according to the target operators (the number of exercises within groups ranged from 4 for *Represent* to 8 for *Order*). Within each group, students were limited in time. If a student did not answer all the exercises from a group, ActiveMath interrupted her/him and transitioned to the next group.

### 3.4 Training Fraction Competencies with Tasks with Typical Errors

In AtuF, tasks with typical errors (TWTW) were used for training fraction competencies. The TWTE were designed around typical errors in the domain identified on the basis of psychological and empirical task and error analyses [22].

We used TWTE instead of conventional fraction tasks to investigate more precisely the question of how efficient the developed feedback strategies are for various typical errors. Addressing this issue with conventional fraction tasks is a methodological challenge due to a high variability of students' error rates and of the types of errors occurring when students work on conventional tasks. Using TWTE for evaluating the effects of the developed feedback components helped us to design and deliver feedback messages that directly target the known error.

Interaction with a TWTE differs from a conventional exercise. In a TWTE students do not commit their own errors by calculating an answer to an exercise; instead, they are provided with a worked step-by-step solution of an exercise that already contains a typical error. Then, as the first task, they are asked to detect the error. If they respond correctly, the system acknowledges it; otherwise, they are provided with a simple knowledge-of-result feedback that indicates the correct option. Figure 1 demonstrates an example of a student interaction with the error-detection phase of a TWTE.

In the second phase, students have to correct the erroneous exercise step (see Figure 2). If they are not able to correct the error on their own, they receive tutoring feedback. Students have three trials to accomplish this task, after each incorrect trial, the feedback details increase. After the third incorrect trial, ActiveMath presents the final feedback in terms of a worked-out solution and moves to the next exercise.

Overall, the treatment session included 30 TWTEs. Students received them in blocks of ten exercises. Within each block, exercises followed the same sequence (in terms of the covered operator): *Represent*, *Represent*, *Transform*, *Transform*, *Add (common)*, *Add (common)*, *Add (uncommon)*, *Add (uncommon)*, *Order*, *Order*. When analysing the data (see Section 4 below), we had to exclude the logs for all TWTEs covering ordering of fractions, because the implementation of these tasks did not follow the same algorithm as for the other eight (instead of separate error-detection and error-correction tasks, students were asked to identify a correct procedure and execute it in one single step). The final set of treatment exercises consists of 24 items.

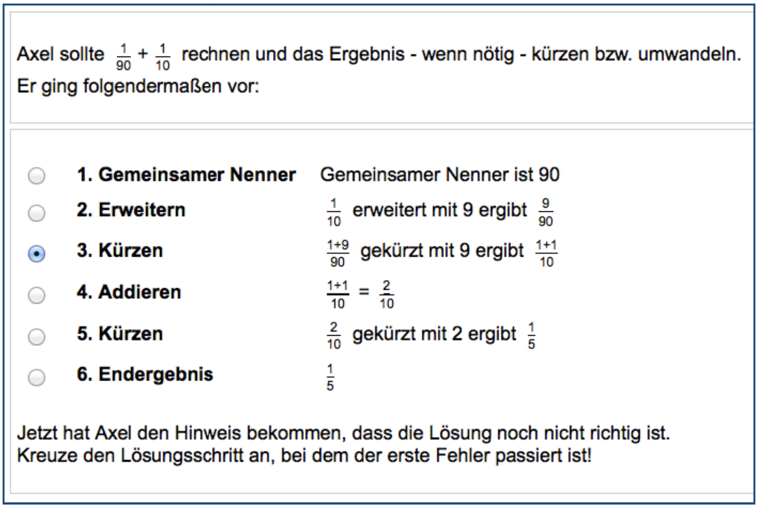



**Fig. 1.** Interface of an AtuF training exercise (error detection task)

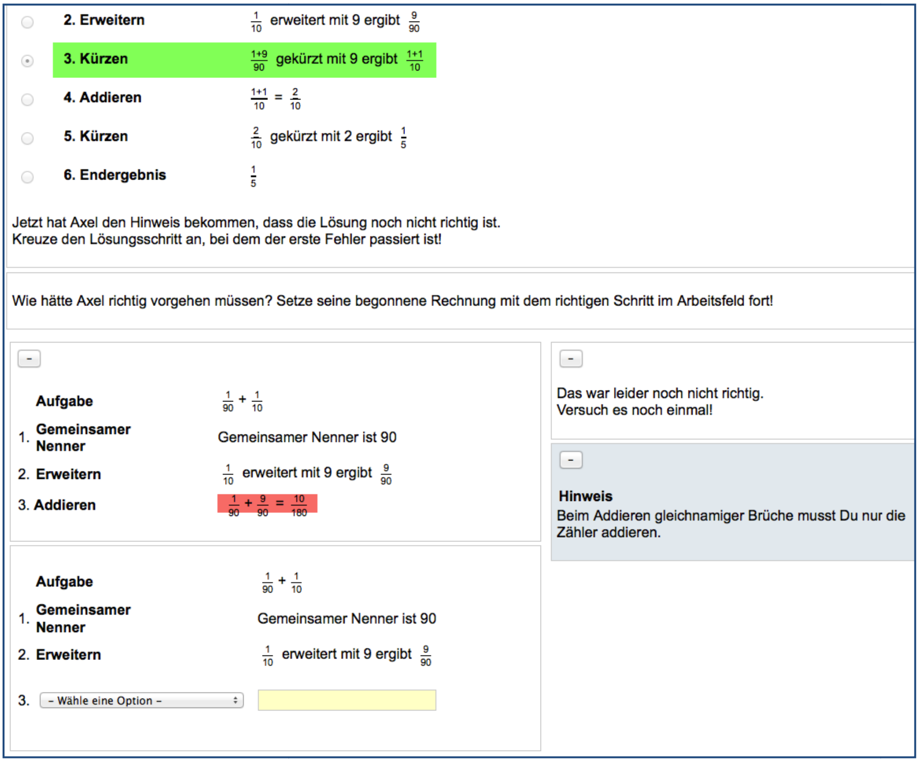



**Fig. 2.** Interface of an AtuF training exercise (error correction task)

### 3.5 Motivation Questionnaire

To assess students' motivation we used the Expectancy-value-Form of domain-specific Learning Motivation (EWF-LM). The EWF-LM has been developed on the basis of an integrative expectancy-value model of motivation (see [15] for more detail), and consists of the four scales: intrinsic value (eight items; e.g., "*I enjoy solving fraction exercises*"; Cronbach's alpha = .87), attainment value (five items; e.g., "*Fraction tasks offer me an exiting occasion for demonstrating my abilities*"; Cronbach's alpha = .73), perceived competence (three items; e.g., "*I am good at solving fraction tasks*"; Cronbach's alpha = .81), and fear of failure (three items; e.g., "*I am worried about mistakes, even if nobody would see them*"; Cronbach's alpha = .81) in relation to fraction tasks. Students had to respond to the EWF-LM items on a 6-point rating scale (0 = "*Not true at all for me*"; 6 = "*Completely true for me*"). Based on the findings from the prior AtuF studies [15, 23], the scales for intrinsic value and perceived competence were selected to guide the feedback adaptation.

### 3.6 Feedback Strategies

Feedback messages in AtuF can vary along two dimensions. From the specificity standpoint, a feedback message can be a hint pointing out an appropriate procedure or a key concept, or it can be a more elaborate explanation describing how the procedure should be applied or why the particular concept is relevant to this task. From the point of type of knowledge communicated by the hint, the feedback can be procedurally or conceptually oriented. For example, in a TWTE on fraction expansion one of the following four messages could be delivered to a student depending on the strategy and the student progress with the task[1]:

- Procedural hint: "When expanding a fraction, one alters the numerator and the denominator equally."
- Procedural explanation: "When expanding a fraction, one alters the numerator and denominator equally. To do this, multiply the numerator and denominator by the same number."
- Conceptual hint: "When expanding a fraction, its value must not change."
- Conceptual explanation: "When expanding a fraction, its value must not change. While the fraction is expanded, the denominator increases, that means the partitioning becomes more fine-grained. But, since the value of the whole fraction does not change, the numerator has to be altered in the same way."

Using the ActiveMath exercise strategy framework described in Section 3.2, several static adaptation strategies for TWTE were developed. They do not alter the error detection phase in any way, but vary presentation of tutoring feedback in the error correction phase of TWTEs. Informed by the results of previous AtuF studies, we have made feedback sequences sensitive to both, student and content characteristics. Table 1 summarizes all feedback conditions.

[1] The messages are translated from German.

**Table 1.** Feedback strategies

| Operator | Strategy A | Strategy B | Strategy C |
|---|---|---|---|
| Represent | *KR; KR; KCR* | *KR; KR; KCR* | *KR; KR; KCR* |
| Transform | *KR; KR+P-Hi; KCR* | *KR; KR+C-Hi; KCR* | *KR; KR; KCR* |
| Add (common) | *KR+P-Hi; KR+P-Ex; KCR* | *KR+C-Hi; KR+C-Ex; KCR* | *KR; KR; KCR* |
| Add (uncommon) | *KR+P-Ex; KR+P-Ex; KCR* | *KR+C-Ex; KR+C-Ex; KCR* | *KR; KR; KCR* |
| Order | *KR+P-Hi; KR+P-Ex; KCR* | *KR+C-Hi; KR+C-Ex; KCR* | *KR; KR; KCR* |

*Note: KR = knowledge of result (correct/incorrect); P-Hi = procedural hint; C-Hi = conceptual hint; P-Ex = procedural explanation; C-Ex = conceptual explanation; KCR = knowledge of the correct response (worked-out solution).*

Three feedback strategies were developed: Strategy A predominantly uses procedural feedback; Strategy B is symmetric to A, but employs conceptual feedback components instead of procedural; Strategy C implements the control condition, it does not use conceptual and procedural hints and explanations, instead it presents simple knowledge-of-result feedback followed by the elaborated worked-out solution. We hypothesized that the procedural strategy A would be more beneficial for students with low motivation, while for highly motivated students either type of feedback strategy might be helpful.

Within strategies A and B, depending on the operator covered by the TWTE, the specificity of feedback differed. Our previous experiments have shown that for some operators higher feedback specificity does not bring any value. For example, exercises related to the *Represent* operator have very high solving rates (>96%) and, thus, do not require elaborated feedback components for the target population.

## 3.7 Design and Procedure

The study consisted of six phases (Figure 3). The pre-questionnaire was used to record demographic data of the students (such as age and gender), as well as to assess their motivational parameters. A short familiarization session was followed by the pre-test assessing students' prior knowledge of fractions. During the training session, students were exposed to different experimental (and control) feedback conditions. Finally, the post-test and post-questionnaires were used to obtain estimations of their posterior knowledge and motivation.

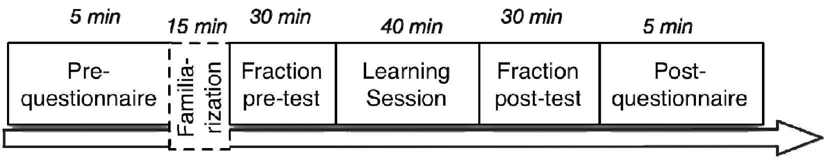



**Fig. 3.** Overview of the experiment procedure

All interactions with the systems were stored in the log files of ActiveMath; this included student answers to questionnaires, assessment and treatment exercises.

Based on the results of the motivation pre-questionnaires, for each student, a cumulative motivation metric was computed as the average score of her answers to the questions about intrinsic motivation and perceived competence. If the metric value was below the threshold = 45 (experimentally identified during the previous AtuF study), the student was categorized as low motivated, otherwise she/he was considered a highly motivated student. After that, ActiveMath automatically assigned the student to one of the three feedback conditions trying to balance the groups with regard to motivation and gender.

## 4 Data Analyses and Results

During the treatment session, students were allowed to work at their own pace with as many tasks as possible within the given time limit. We inspected ActiveMath log files to identify the data of students who (a) worked too slowly to complete even the first block of 8 TWTEs, (b) succeeded in correcting all errors without any support (i.e., were not provided with tutoring feedback at all). We excluded these cases from further data analyses, since tutoring feedback messages can only be beneficial if students observe them at least several times [12]. We also had to exclude a few incomplete logs (data was missing due to system failures). The resulting sample used in this study consists of data from 128 students (57 female, 66 male; five preferred not to answer).

Table 2 shows how the students with different motivation profiles (i.e., low perceived competence and intrinsic value vs. high perceived competence and intrinsic value) have been assigned to the feedback conditions.

**Table 2.** Cross tabulation of feedback strategies and students' motivation

| | | perceived competence & intrinsic value | | |
|---|---|---|---|---|
| | | low | high | Total |
| Feedback Strategy | procedural | 17 | 36 | 53 |
| | conceptual | 19 | 36 | 55 |
| | control | 8 | 12 | 20 |
| Total | | 45 | 85 | 128 |

### 4.1 Preliminary Analyses

Prior to performing statistical data analyses, we screened the data for outliers and normality. While performing the analyses, assumptions of the procedures were tested and met, unless otherwise noted. Furthermore, since the assignment to the feedback strategies was done based on students' level of perceived competence and intrinsic value, we checked if there were any differences in the level of fraction competency prior to the treatment (performance on the pre-test) and for the first TWTE of each type during the treatment (baseline of successful error corrections in the treatment;

see Table 3 for descriptive data). The analysis revealed no significant differences for these variables across different feedback strategies (all F<1), yet the low motivated students under the procedural feedback condition started with a significantly lower level of fraction competency as demonstrated by the significant interaction among the factors *feedback* and *motivation* ($F(2,122) = 3.54; p=.03; eta^2=.06$). These students had also the lowest baseline regarding their success-rate in correcting the TWTE, yet the difference there was not statistically significant.

**Table 3.** Overview of means and standard deviations for competency levels, number of TWTE worked on, baseline of successful error correction in the treatment and feedback efficiency by feedback conditions and motivation

| | | | | Feedback strategy | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Motivation | | | procedural | | conceptual | | control | |
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Competency level Pre-test | low | 14.61 | 5.38 | 12.65 | 6.20 | 15.74 | 4.75 | 16.13 | 4.02 |
| | high | 16.38 | 4.40 | 17.28 | 4.79 | 15.58 | 4.02 | 16.08 | 4.14 |
| Competency level Post-test | low | 16.59 | 4.75 | 16.59 | 4.37 | 16.11 | 5.03 | 17.75 | 5.26 |
| | high | 18.32 | 3.97 | 18.89 | 4.58 | 17.89 | 3.54 | 17.92 | 3.23 |
| Number of TWTE worked on | low | 19.00 | 4.46 | 18.47 | 4.74 | 19.58 | 3.67 | 18.75 | 5.87 |
| | high | 16.59 | 3.95 | 16.47 | 3.73 | 16.39 | 4.14 | 17.58 | 4.23 |
| Baseline Treatment | low | 2.09 | 1.09 | 2.00 | 1.17 | 2.05 | 1.13 | 2.38 | .92 |
| | high | 2.21 | 1.05 | 2.31 | 1.04 | 2.14 | 1.13 | 2.17 | .94 |
| Feedback efficiency step 1 | low | .25 | .28 | .40 | .31 | .16 | .25 | .18 | .19 |
| | high | .27 | .24 | .27 | .24 | .25 | .20 | .32 | .34 |
| Overall feedback efficiency | low | .33 | .30 | .45 | .32 | .27 | .30 | .22 | .17 |
| | high | .37 | .28 | .42 | .27 | .32 | .23 | .39 | .42 |

### 4.2 Differential Effects of Feedback during Treatment

During the treatment, the participants had up to three attempts per error-correction task. If a student failed on the first attempt, she received a feedback message addressing this error according to the experimental condition. To verify how effective these messages were in terms of helping students correct the typical error in the next attempt with the task, we have computed two measures characterizing feedback efficiency: *initial feedback efficiency* and *overall feedback efficiency* (see Table 3). The first measure quantifies the efficiency of feedback provided after the first unsuccessful error correction attempt, and is computed as the number of successful error corrections after the first feedback divided by the difference between the number of all attempted tasks and the number of tasks solved correctly on the first attempt (i.e., tasks solved without any feedback). Additionally, the overall feedback efficiency is computed as the total number of successful error correction attempts after feedback of any level divided by the difference between the number of all attempted tasks and the number of tasks solved correctly on the first attempt (before seeing any feedback).

A *feedback* (3) by *motivation* (2) ANOVA for *initial feedback efficiency* revealed a significant main effect of feedback, $F(2,122) = 3.34; p=.04; eta^2=.05,$ and a two-way interaction $F(2,122) = 3.05; p=.05; eta^2=.05$. For students with low motivation, procedural feedback strategy was significantly more efficient than the conceptual feedback strategy or the control condition strategy that provided a worked-out solution after the third incorrect attempt. For students with high motivation, there were no statistically significant differences in feedback efficiency after the first incorrect attempt (see Figure 4).

Regarding the overall feedback efficiency, we have also found a significant main effect of feedback $F(2,122) = 3.14; p=.05; eta^2=.05$. The procedural feedback strategy was more efficient than the conceptual and the control strategy (see Figure 4).

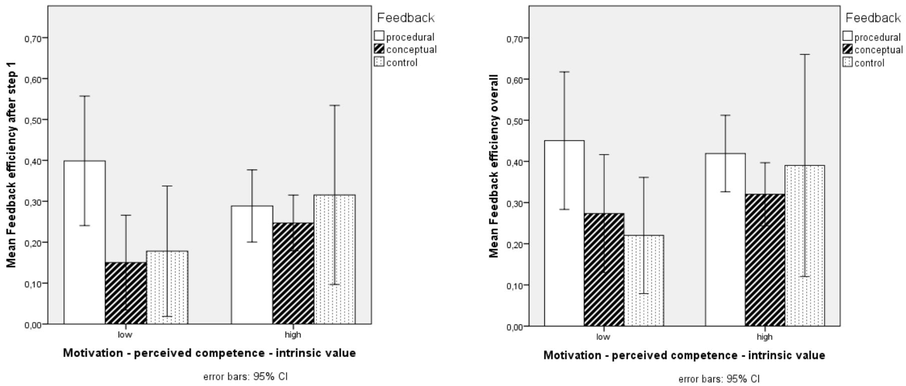



**Fig. 4.** Feedback efficiency after receiving feedback on the first unsuccessful error-correction attempt (left), and over all attempts (right) for students with low and high motivation by feedback condition

## 4.3 Differential Effects of Feedback on Learning Gain

The role of feedback as a teaching intervention is not only to facilitate problem solving, but, first and foremost, to improve students' learning. To investigate if, how and under which individual conditions, the feedback strategies affected knowledge gain from pre- to post-test, we run a repeated-measure ANOVA with the between factors feedback (3), motivation (2), and the within factor (2) for the fraction competency measures of the pre- and post-test.

This analysis revealed a significant three-way interaction with *Wilks' Lambda* = *.94*, $F(2, 122) = 4.07$, $p = .02$, $eta^2 = .06$. Students with low perceived competence and intrinsic value improved their level of fraction competency significantly more with the procedural feedback strategy than with the conceptual or the control strategy. It is noteworthy, that under conceptual feedback condition, they had the smallest knowledge gain. Students with high perceived competence and intrinsic value improved their fraction competency level comparably under all feedback strategies. Figure 5 illustrates these findings.

It should be noted that low motivated students with the procedural feedback strategies had a significantly lower level of prior competency in the pre-test than all other groups. Yet, for all groups there was room for improvement since the fraction tests consisted of 29 items. Moreover, it seems interesting that irrespective of the feedback strategy, the low motivated students worked on significantly more TWTEs during treatment than the highly motivated students ($F(2,122) = 6.17$; $p=.01$; $eta^2=.05$), (see Table 3 for the descriptive data).

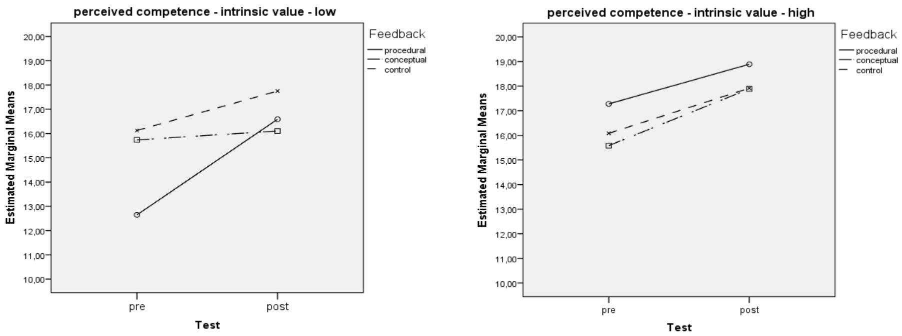



**Fig. 5.** Learning gain from pre- to post-test for students with low vs. high perceived competence and intrinsic value by feedback condition

## 5 Summary and Conclusions

This paper investigates the issue of how tutoring feedback within an AES can be adapted to the pre-existing level of student motivation. The domain of basic fraction arithmetic was used as an example; and the study focused on the two motivation parameters: perceived competence and intrinsic motivation. The findings support the hypothesis that students with low motivation would benefit from a feedback strategy providing procedural tutoring feedback messages more than from a feedback strategy with conceptual feedback messages. The findings also show that, for students with high motivation, either type of feedback strategy works equally well.

This experiment reveals some challenges of applying the ITF-model to the evaluation of tutoring feedback strategies. In order to provide students with occasions to use the feedback information for correcting the errors of the TWTE, we implemented two static multi-trial feedback strategies. These feedback strategies provided students with up to three trials to correct errors, and students were free to work at their own pace. The next level of feedback was offered only when a student did not succeed on the previous trial with the less detailed level of feedback. Through this approach, we obtained a rich dataset for investigating how feedback conditions, feedback properties and student feedback processing contribute to feedback efficiency. Yet, we were faced with a number of methodological challenges, including high student variability in terms of working pace, and the number of trials needed to complete tasks. The findings indicate, for example, that students with low motivation worked on more TWTEs than students with high motivation, but they had a lower success rate in completing the error correction step of TWTEs. Further analysis is needed to more accurately account for the actual feedback processing. We plan to take a deeper look at the the timestamp data of the log files in order to investigate if and how the time students spend on answering the task after a feedback message is connected to successful error correction.

Another challenge relates to the issue of how to assign students to the feedback strategies. Based on our prior findings, the students of the present study were assigned to the two feedback strategies according to their pre-existing level of motivation. Their initial level of competency was controlled, but not used as an assignment criterion. To classify students' initial levels of motivation, the assignment procedure used the median of the motivational variables obtained in a prior study with 207 students from the same population (6$^{th}$ and 7$^{th}$ graders from Dresden schools) [15]. This assignment procedure proved to have limitations, since the median was different for the present sample compared to the sample of the prior study. Furthermore, the competency level of the three low motivation groups was rather dissimilar on the pre-test. Further development and evaluation of the classification and assignment procedure integrating motivational, cognitive, and metacognitive variables into a feedback adaptation strategy is an interesting practical and methodological challenge that we plan to address in our future work.

## References


1. Hattie, J., Timperley, H.: The power of feedback. Review of Educational Research 77, 81–112 (2007)
2. Hattie, J., Gan, M.: Instruction based on Feedback. In: Mayer, R., Alexander, P. (eds.) Handbook of Research on Learning and Instruction, pp. 249–271. Routledge, New York (2011)
3. Narciss, S.: Feedback strategies for interactive learning tasks. In: Spector, J.M., Merrill, M.D., van Merrienboer, J.J.G., Driscoll, M.P. (eds.) Handbook of Research on Educational Communications and Technology, pp. 125–144. Lawrence Erlbaum Associates, Mahwah (2008)
4. Shute, V.J.: Focus on formative feedback. Review of Educational Research 78, 153–189 (2008)
5. Ramaprasad, A.: On the definition of feedback. Behavioral Science (28), 4–13 (1983)
6. Sadler, D.R.: Formative assessment and the design of instructional systems. Instructional Science 18, 119–144 (1989)
7. Corbalan, G., Paas, F., Cuypers, H.: Computer-based feedback in linear algebra: Effects on transfer performance and motivation. Computers & Education 55, 692–703 (2010)

8. Mitrovic, A., Ohlsson, S., Barrow, D.: The effect of positive feedback in a constraint-based intelligent tutoring system. Computers & Education 60, 264–272 (2013)
9. Dennis, M., Masthoff, J., Mellish, C.: User modeling, adaptation, and personalization. In: Masthoff, J., Mobasher, B., Desmarais, M.C., Nkambou, R. (eds.) UMAP 2012. LNCS, vol. 7379, pp. 297–302. Springer, Heidelberg (2012)
10. Conati, C., Manske, M.: Evaluating adaptive feedback in an educational computer game. In: Ruttkay, Z., Kipp, M., Nijholt, A., Vilhjálmsson, H.H. (eds.) IVA 2009. LNCS (LNAI), vol. 5773, pp. 146–158. Springer, Heidelberg (2009)
11. Goldin, I., Koedinger, K., Aleven, V.: Learner differences in hint processing. In: Yacef, O.Z.K., Hershkovitz, H., Yudelson, M., Stamper, J. (eds.) Proceedings of the 5th International Conference on Educational Data Mining, pp. 73–80. International EDM Society, Worcester (2012)
12. Narciss, S.: Designing and Evaluating Tutoring Feedback Strategies for digital learning environments on the basis of the Interactive Tutoring Feedback Model. Digital Education Review 23, 7–26 (2013), `http://greav.ub.edu/der`
13. Butler, D., Winne, P.: Feedback and self-regulated learning: A theoretical synthesis. Review of Educational Research 65, 245–281 (1995)
14. Klieme, E., Leutner, D., Kenk, M.: Kompetenzmodellierung. Zeitschrift für Pädagogik 56, 9–11 (2010)
15. Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., Melis, E.: Exploring feedback and student characteristics relevant for personalizing feedback strategies. Computers and Education 71, 56–76 (2014)
16. Eichelmann, A., Narciss, S., Faulhaber, A., Melis, E.: Analyzing computer-based fraction tasks on the basis of a two-dimensional view of mathematics competencies. In: Zumbach, J., et al. (eds.) Beyond Knowledge: The Legacy of Competence. Meaningful Computer-based Learning Environments, pp. 125–134. Springer, Dordrecht (2008)
17. Melis, E., Goguadze, G., Homik, M., Libbrecht, P., Ullrich, C., Winterstein, S.: Semantic-aware components and services of ActiveMath. British Journal of Educational Technology 37(3), 405–423 (2006)
18. Sosnovsky, S., Dietrich, M., Andrès, E., Goguadze, G., Winterstein, S., Libbrecht, P., Siekmann, J., Melis, E.: Math-Bridge: Closing the Gaps in European Remedial Mathematics with Technology-Enhanced Learning. In: Wassong, T., et al. (eds.) Mit Werkzeugen Mathematik und Stochastik lernen – Using Tools for Learning Mathematics and Statistics, pp. 437–451. Springer, Heidelberg (2014)
19. Goguadze, G.: Representation for interactive exercises. In: Carette, J., Dixon, L., Coen, C.S., Watt, S.M. (eds.) Calculemus/MKM 2009. LNCS (LNAI), vol. 5625, pp. 294–309. Springer, Heidelberg (2009)
20. Goguadze, G., Melis, E.: One Exercise - Various Tutorial Strategies. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (eds.) ITS 2008. LNCS, vol. 5091, pp. 755–757. Springer, Heidelberg (2008)
21. Schnaubert, L.: Kompetenzdiagnostik in der Mathematik: Konstruktion und empirische Evaluation von Bruchadditionsaufgaben auf Grundlage eines zwei-dimensionalen Kompetenzmodells, Fachrichtung Psychologie. Technische Universität Dresden (2010)
22. Eichelmann, A., Narciss, S., Schnaubert, L., Melis, E.: Typische Fehler bei der Addition und Subtraktion von Brüchen – Ein Review zu empirischen Fehleranalysen. Journal für Mathematik-Didaktik 33(1), 29–57 (2012)
23. Schnaubert, L., Andrès, E., Narciss, S., Eichelmann, A., Goguadze, G., Melis, E.: Student Behavior in Error-Correction-Tasks and Its Relation to Perception of Competence. In: Kloos, C.D., Gillet, D., Crespo García, R.M., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 370–383. Springer, Heidelberg (2011)

# Learning Resource Recommendation: An Orchestration of Content-Based Filtering, Word Semantic Similarity and Page Ranking

Nguyen Ngoc Chan, Azim Roussanaly, and Anne Boyer

Université de Lorraine
LORIA UMR 7503, France
{ngoc-chan.nguyen,azim.roussanaly,anne.boyer}@loria.fr

**Abstract.** Technologies supporting online education have been abundantly developed recent years. Many repositories of digital learning resources have been set up and many recommendation approaches have been proposed to facilitate the consummation of learning resources. In this paper, we present an approach that combines three recommendation technologies: content-based filtering, word semantic similarity and page ranking to make resource recommendations. Content-based filtering is applied to filter syntactically learning resources that are similar to user profile. Word semantic similarity is applied to consolidate the content-based filtering with word semantic meanings. Page ranking is applied to identify the importance of each resource according to its relations to others. Finally, a hybrid approach that orchestrates these techniques has been proposed. We performed several experiments on a public learning resource dataset. Results on similarity values, coverage of recommendations and computation time show that our approach is feasible.

**Keywords:** Learning Resource, Technology Enhanced Learning, Content-based Filtering, Word Semantic Similarity, TF-IDF, VSM, PageRank.

## 1 Introduction

During the past few years, along with the development of technologies supporting online education, numerous repositories of digital learning resources have been set up, such as MERLOT[1], OER Commons[2] and LRE For Schools[3] [21]. They provide open learning resources of various disciplines (such as arts, humanities, science and technologies, etc.), levels (such as primary school, secondary school, high school, higher education, etc.) and types (such as lab, lecture note, exercise, tutorial, etc.). These resources allow users to self-study or to consolidate their knowledge on different domains. However, the variety of these resources, in contrast, easily discourage users to continue studying. For instance, whenever users want to study a subject, they search or browse resources related to that subject

[1] http://www.merlot.org
[2] http://www.oercommons.org
[3] http://lreforschools.eun.org

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 302–316, 2014.
© Springer International Publishing Switzerland 2014

and preview them using try-and-error method. They spend much time to reach to their expected resources. In addition, after learning a resource, they should redo the search/browse process if they want to find other related resources.

In order to encourage the usage of online learning resources, recommender systems are considered as a pivotal solution, especially on the Technology Enhanced Learning (TEL) domain [14]. Many approaches that apply recommendation techniques to support resource recommendation have been proposed. For example, they apply collaborative filtering [12,20], content-based filtering [9,11], examine user ratings [5,15], study association rules [12,19] or analyze user feedback [8]. Bayesian model [1], Markov chain [6], resource ontologies [16,19] and hybrid models [16,7] were also proposed. However, most of existing systems still remain at a design or prototyping stage. Only few of them have been reported to be evaluated through trials that involved human users [14].

In our work, we also target to encourage the usage of online learning resources with recommendations. However, different from existing approaches, we propose an innovative solution that orchestrate 3 recommendation techniques: *content-based filtering*, *word semantic similarity* and *page ranking*. Content-based filtering is applied to filter syntactically learning resources that are similar to user profile. Word semantic similarity is applied to consolidate the content-based filtering with word semantic meanings. Page ranking, which is inherited from the Google PageRank algorithm [2], is applied to identify the importance of each resource according to its relations to others. By hybridizing these techniques, our objective is three-fold: (i) to present an important application of recommendation techniques on a specific domain, which is online education, (ii) to show a possible combination of existing techniques to develop a hybrid recommendation approach, and (iii) to demonstrate a retrieval of important items that are not only syntactically but also semantically relevant to a request.

As keywords present concisely and precisely the content of resource, whereas recent viewed resources present recent user interest, we propose to build implicitly user profile based on keywords of recent viewed resources. By building user profile based on historical keywords, our approach is able to make recommendations that are close to user recent interest. In addition, it does not ask any effort from users such as completing registration form, specifying preferences, etc.

The paper is organized as follows: the next section elaborates in detail recommendation techniques applied in our approach and their combination. Experiments are presented in section 3. Related word is discussed in section 4 and we conclude our approach in section 5.

## 2 Learning Resource Recommendation

In this section, we present in detail approach to recommend learning resources for an active user. We firstly introduce a basic CB filtering approach that applies vector space model (section 2.1). Then, we show a refinement of text vectors based on word semantic similarity (section 2.2) and the ranking of resources based on their relations (section 2.3). Finally, we present a hybrid approach that combines these techniques (section 2.4).

### 2.1 Content-Based Filtering with Vector Space Model

Vector Space Model (VSM) is a method popularly used in Information Retrieval to compute the similarity between documents. It considers each word as a dimension. It presents each document as a vector in the $n$-dimensional space of words. Elements of each vector are weights of the corresponding words in the document, which present their importance in the corpus. These weights are computed by the term-frequency (TF) and inverse-document-frequency (IDF). Concretely, consider a documents $d_i$, which is presented by a vector $\overrightarrow{d_i} = \{w_{i1}, w_{i2}, \ldots, w_{in}\}$, where $n$ is the number of words in the corpus and $w_{ik}$ $(k = \overline{1..n})$ is the weight of the $k^{th}$ element in the vector. $w_{ik}$ is computed by Eq. 1

$$w_{i,k} = TF(i,k) \times IDF(k) = \frac{|w_k|}{|d_i|} \times \log \frac{n}{|D_k|} \tag{1}$$

where $|w_k|$ is the occurrence of the word $w_k$ in $d_i$, $|d_i|$ is the number of words in $d_i$, and $|D_k|$ is the number of documents containing $w_k$.

Then, similarity between documents is computed by the cosine of the angle between their representative vectors. For example, similarity between two documents $d_i$ and $d_j$ is computed by Eq. 2.

$$sim(d_i, d_j) = cosine(\overrightarrow{d_i}, \overrightarrow{d_j}) = \frac{\overrightarrow{d_i} \cdot \overrightarrow{d_j}}{|\overrightarrow{d_i}| \times |\overrightarrow{d_j}|} \tag{2}$$

In our approach, we apply VSM to compute the similarity between a user profile and a resource description. User profile is defined by the set of keywords of his recent viewed resources, whereas resource description is all the text that are used to describe a resource[4].

Concretely, consider a corpus that consists of $m$ learning resources and an active user $u_a$ is viewing a resource $r_a$. Assume that $m$ resources contain $n$ different words. Let $d_i$ be text description of resource $r_i$ $(i = \overline{1..m})$. We present $d_i$ as a vector $\overrightarrow{d_i} = \{w_{i1}, w_{i2}, \ldots, w_{in}\}$. $w_{ik}$ $(k = \overline{1..n})$ is the TF-IDF weight (computed by Eq. 1) of the $k^{th}$ corresponding word in the resource description.

Let $\{k_1, k_2, \ldots, k_t\}$ be $t$ keywords of $h$ recent viewed resources of $u_a$. We consider these historical keywords as a query $q_a$. We present $q_a$ as a vector $\overrightarrow{q_a} = \{w_{a1}, w_{a2}, \ldots, w_{an}\}$ in the same space with resource descriptions. Elements of $\overrightarrow{q_a}$ are TF-IDF weights of corresponding words in the query.

Then, according to Eq. 2, similarity between $q_a$ and $d_i$ is given by Eq. 3.

$$sim(q_a, d_i) = \frac{\overrightarrow{q_a} \cdot \overrightarrow{d_i}}{|\overrightarrow{q_a}| \times |\overrightarrow{d_i}|} = \frac{\sum_{k=1}^{n} w_{ak} w_{ik}}{\sqrt{\sum_{k=1}^{n} w_{ak}^2} \times \sqrt{\sum_{k=1}^{n} w_{ik}^2}} \tag{3}$$

[4] In our experiments, $d_i$ includes words in the title, abstract, keywords, discipline and classification of the resource.

We apply Eq. 3 for all $d_i \in \{1, 2, \ldots, N\}$. Then, we sort resources in descending order according to the computed similarity with $q_a$. Finally, top-K resources are selected as the relevant resources of the active user's historical keywords.

As words can appear in different forms (singular, plural) and tenses (present, past, future), we proceed words stemming and stop-words removing before performing the similarity computation. In addition, as keywords can be singular words or compound words, we preprocess resource descriptions by identifying compound words that are matched to those in the historical keywords. By this identification, we can treat compound words as singular words.

For example, consider a query and a resource $r_i$ with description $d_i$ as follows:

$q_a$ ={recommender system, technology enhanced learning, learning resources}
$d_i$ ={On the technology enhanced learning domain, recommender systems are consider as a pivotal solution to recommend learning resources.}

After words stemming and stop-words removing, singular word treatment identifies the word occurrence in $q_a$ and $d_i$ as follows.

$q_a$: {recommend(1), system(1), technology(1), enhance(1), learn(2), resource(1)}

$d_i$: {technology(1), enhance(1), learn(2), domain(1), recommend(2), system(1), consider(1), pivot(1), solution(1), resource(1) }

meanwhile, compound word treatment identifies the word occurrence in $q_a$ and $d_i$ as follows.

$q_a$: {recommend system (1), technology enhance learn (1), learn resource (1)}

$d_i$: {technology enhance learn (1), domain(1), recommend system (1), consider(1), pivot(1), solution(1), recommend (1), learn resource (1)}

Based on these word occurrences, TF-IDF (Eq. 1) and VSM (Eq. 3) are applied to computed the similarity between $q_a$ and $d_i$.

### 2.2 Query-Resource Matching Based on Word Semantic Similarity

Polysemy and synonymy are common problems facing in text processing. If we deal with only word syntactical matching, without considering the semantic similarity, we easily miss potential matchings of different words which expose the same meaning. In this section, we present an integration of word semantic similarity in our approach in order to recommend more precisely learning resources.

There exist many research on the word semantic similarity that can be applied in our approach, such as [10,4,13,18]. However, as we focus on the matching between resources instead of semantic similarity, discussion about this topic is out of scope of our paper. In our experiment, we adopt the work of Peter Kolb [10], which is a high accurate approach based on the co-occurrence of words in the Wikipedia dataset, to compute the word semantic similarity.

Consider an active user $u_a$ who has recently viewed $h$ resources which have a list of keywords $q_a = \{k_1, k_2, \ldots, k_t\}$. We consider this list as a query. For each resource $r_i$, which is considered to match with $q_a$, we propose to replace each word in the resource description, i.e $d_i$, by its most *semantically* similar word in the query if this word does not appear in the query. We update the weight of words in the resource description according to their semantical similarity with the selected words in the query . Finally, we weight words in both resources and query by TF-IDF and compute their similarity by applying VSM.

Concretely, consider a word $v_x$ in resource description $d_i$ with an occurrence $o_x$. Suppose that $v_x$ is most semantically similar to a word $k_y$ in $q_a$ with a similarity value $s(v_x, k_y) \in (0, 1)$. We substitute $v_x$ in $d_i$ by $k_y$ and update its weight to $w_{xy} = o_x s(v_x, k_y)$. This means that $o_x$ times that $v_x$ appears in $d_i$ is considered as $o_x s(v_x, k_y)$ times that $k_y$ appears in $d_i$. We repeat this substitution for all words in $d_i$.

Recall the example of $q_a$ and $d_i$ in section 2.1. The substitution of words in $d_i$ by the most semantically similar words in $q_a$ is given in Table 1. For instance, the word 'domain' in $d_i$ is the most similar to the word 'resource' in $q_a$ (similarity=0.015), we replace the word 'domain' by 'resource' and update its weight to $1 \times 0.015 = 0.015$ and so on.

**Table 1.** Example of word substitution based on semantic similarity

| | Similarity with words in $q_a$ | | | | | | Substitution | |
|---|---|---|---|---|---|---|---|---|
| $v_x$ | recommend | system | technology | enhance | learn | resource | $k_y$ | $w_{xy}$ |
| domain | 0.002 | 0.007 | 0.009 | 0.005 | 0.001 | **0.015** | resource | 0.015 |
| consider | **0.056** | 0 | 0.002 | 0.004 | 0.035 | 0.003 | recommend | 0.056 |
| pivot | 0.001 | 0.003 | 0.002 | **0.0043** | 0.001 | 0.0039 | enhance | 0.0043 |
| solution | 0.003 | 0.015 | **0.017** | 0.005 | 0.001 | 0.014 | technology | 0.017 |

Assume that $n_1$ words in $d_i$ are replaced by $k_1$ with the updated weights are $\{w_{11}, w_{21}, \ldots, w_{n_1 1}\}$, $n_2$ words in $d_i$ are replaced by $k_2$ with updated weights are $\{w_{12}, w_{22}, \ldots, w_{n_2 2}\}$, and so on, and $n_0$ words in $d_i$ are not replace by any word in the query.

The resource description $d_i$ becomes $d'_i = \{k_1, k_2, \ldots, k_t, k_{t+1}, \ldots, k_{t+n_0}\}$ with corresponding weights are $\{\sum_{j=1}^{n_1} w_{j1}, \sum_{j=1}^{n_2} w_{j2}, \ldots, \sum_{j=1}^{n_k} w_{jk}, 0, \ldots, 0\}$.

Similarity between $l_a$ and $d_i$ is calculated by similarity between $q_a$ and $d'_i$. As weights of $k_{t+1}, \ldots, k_{t+n_0}$ are 0, we can remove them in $d'_i$. $d'_i$ becomes a vector in the same space with $q_a$. We, then, compute TF-IDF (Eq. 1) for words in both $q_a$ and $d_i$. Finally, we apply VSM (Eq. 3) to calculate their similarity.

For example, after substitute words in $d_i$ according to Table 1, we obtain $d'_i$={recommend(2.056), system(1), technology(1.017), enhance(1.0043), learn(2), resource(1.015)}, which is a vector in the same space with $q_a$. Then, we can apply TF-IDF and VSM to compute their similarity.

According to similarity between the query and all resources, we make recommendations by selecting the top-K most similar resources. In experiment, we run our approach and compare results in two cases: recommendations with and without word semantic similarity.

### 2.3 Resource Ranking Inspired from Google PageRank Algorithm

The importance of a resource in a corpus can be evaluated by different criteria such as the knowledge provided by that resource, its applications on different domains, its relations to other resources, or just the number of users viewing that resource. In this section, we present a ranking algorithm to evaluate the importance of resources based on their relations. This algorithm is inspired from the Google PageRank algorithm [2].

According to [2], rank of a page $A$ is computed by Eq. 4.

$$PR(A) = (1 - d) + d \times \sum_{i=1}^{n} \frac{PR(T_i)}{C(T_i)} \tag{4}$$

where $0 \leq d \leq 1$ is a damping factor (in [2], $d = 0.85$), $T_1, T_2, \ldots, T_n$ are pages which point to $A$ and $C(T_i)$ is the number of links going out of $T_i$. Initial page rank of each page is $\frac{1}{N}$ where $N$ is the number of pages. Page ranks of all pages are iteratively updated by Eq. 4 until they achieve stable values within a given threshold.

By another point of view[17,22], page ranks are defined as a vector $v^*$ that satisfies:

$$Gv^* = v^* \tag{5}$$

where $G$ is the Google matrix, which is defined as:

$$G = \frac{1 - d}{N} S + dM \tag{6}$$

where $S$ is the matrix with all entries equal to 1 and $M$ is a transition matrix.

The transition matrix $M$ presents links between pages. Value of an element $M_{[i,j]}$ is the weight of the link from page $j^{th}$ to page $i^{th}$. $M_{[i,j]}$ satisfies $\sum_{i=1}^{N} M_{[i,j]} = 1, \forall j = \overline{1..N}$. According to [2], $M$ is a Markov matrix and if a page $j$ has $k$ out-going links, each of them a weight $\frac{1}{k}$.

According to Eq. 5, $v^*$ is the eigenvector of the Markov matrix $G$ with the eigenvalue 1. Let $v_0$ be the initial page rank vector, elements in $v_0$ are set to $\frac{1}{N}$. $v^*$ is iteratively computed as following:

$$v_{i+1} = Gv_i \tag{7}$$

until $|v_{i+1} - v_i| < \epsilon$ ($\epsilon$ is a given threshold).

As $G$ is a Markov matrix, $v_{i+1}$ will converge to $v^*$ after certain iterations. $v^*$ presents the ranking of web pages according to their hyperlink.

Inspired by the Google PageRank algorithm, we propose an algorithm to compute the ranking of learning resources. In our algorithm, we take into account resource relations instead of hyperlink between pages.

Basically, a resource can be a part of another resource, include other resources or associate to other resources. We define these relations respectively *'is part of'*, *'contains'* and *'associates to'*, in which 'is part of' is a 1-1 relation, 'contains' is a 1-n relation and 'associates to' is an n-n relation (Fig. 1).

Each relation not only presents a hyperlink between resources, but also exposes a particular semantic meaning. For instance, 'associates to' indicates a set of coherent resources that supplement each other to present some knowledge;

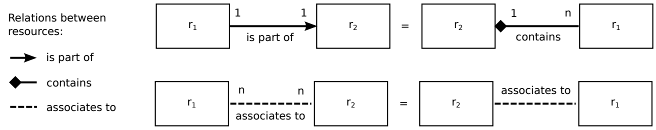



**Fig. 1.** Relations between resources

'contains' lists out a set of resources to be involved within a subject and some of them could be not coherent; 'is part of' signifies a resource which is a member of another resource but does not clearly present involved related resources. According to these meanings, we propose to assign a relation weight to each relation type. Concretely, 'associated to' has more weight than 'contains' and 'contains' has more weight than 'is part of'.

Let $w_{ra}$, $w_{rc}$ and $w_{rp}$ be weights of 'associates to', 'contains' and 'is part of', we have $w_{ra} > w_{rc} > w_{rp}$. For simplicity, we set[5]:

$$w_{ra} = \alpha w_{rc} = \alpha^2 w_{rp}, 0 < \alpha \leq 1 \tag{8}$$

In Google PageRank algorithm, weights of all hyperlink are set to be equal. In our approach, we weight relations between resources according to their types instead of giving an average weight for all relations. Concretely, assume that a resource $r_i$ has $a$ 'associates to' relations, $b$ 'contains' relations and $c$ 'is parts of' relations, we have:

$$aw_{ra} + bw_{rc} + cw_{rp} = 1 \tag{9}$$

From Eq. 8 and Eq. 9, we have:

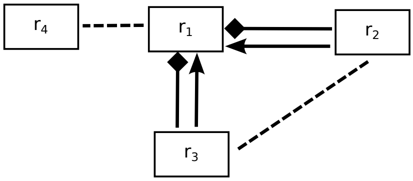
$$w_{rp} = \frac{1}{a\alpha^2 + b\alpha + c}; \quad w_{rc} = \frac{\alpha}{a\alpha^2 + b\alpha + c}; \quad w_{ra} = \frac{\alpha^2}{a\alpha^2 + b\alpha + c} \tag{10}$$

Eq. 9 ensures that the matrix $M$ and $G$ in Eq. 6 are Markov matrices. Hence, the multiplication of these matrices with an initial weighted-vector will converge to an eigenvector vector. It means that we can compute the PageRank vectors of resources based on the new weights. The weights calculated by Eq. 10 are used to initialize the matrix $M$ and $G$ in Eq. 6. Then the ranking vector $v^*$ is calculated by Eq. 7.

For example, Fig 2 presents relations of 4 resources ($r_1$, $r_2$, $r_3$, $r_4$) in a corpus. Based on these relations, we can compute the weights of each relations ($w_{ra}$, $w_{rc}$, $w_{rp}$) and create the relation matrix $M$ (the right column in Fig 2). Each element $M_{[i,j]}$ presents the weight of relation from $r_j$ to $r_i$ and the matrix satisfy that sum of all elements in each column is equal to 1. Based on $M$, we can easily compute $G$ and $v^*$ by applying Eq. 6 and Eq. 7.

The PageRank vector presents the importance of resources according to their relations. Their rankings can be used in a resource searching application, similar

[5] In our experiment, we set $\alpha = 0.9$.




$r_1$: $a = 1, b = 2, c = 0,$
$r_2$: $a = 1, b = 0, c = 1,$
$r_3$: $a = 1, b = 0, c = 1,$
$r_4$: $a = 1, b = 0, c = 0,$

With $\alpha = 0.9$, we have:
$r_1$: $w_{ra} = 0.31, w_{rc} = 0.345;$
$r_2$: $w_{ra} = 0.45, w_{rp} = 0.55;$
$r_3$: $w_{ra} = 0.45, w_{rp} = 0.55;$
$r_4$: $w_{ra} = 1;$

and the matrix $M$:

| | $r_1$ | $r_2$ | $r_3$ | $r_4$ |
|---|---|---|---|---|
| $r_1$ | 0 | 0.55 | 0.55 | 1 |
| $r_2$ | 0.345 | 0 | 0.45 | 0 |
| $r_3$ | 0.345 | 0.45 | 0 | 0 |
| $r_4$ | 0.31 | 0 | 0 | 0 |

**Fig. 2.** Example of resource relations and the corresponding matrix

to page rankings are used in the Google search engine. They can also be considered as a parameter in a recommendation application in order to refine the recommendation result.

### 2.4 A Hybrid Recommendation Approach

The content-based filtering algorithm (section 2.1) and its refinement with word semantic similarity (section 2.2) enable the retrieval of resources that are syntactically and semantically similar to user profile. Meanwhiles, resource ranking (section 2.3) helps to identify the importance of resources based on their relations. Therefore, their combination possibly retrieves resources that are both important and relevant to a user.

In our approach, we propose to multiply the similarity between user profile and resources with the ranking of resources to infer their final matching scores. Concretely, consider an active user $u_a$ who has a profile $q_a$ and a resource $r_i$ which has a description $d_i$. Let $d'_i$ be the refined resource of $d_i$ by applying word semantic similarity according to the query $q_a$. The final matching score between $q_a$ and $r_i$ is given by Eq. 11.

$$scr(q_a, r_i) = sim(q_a, d'_i) \times v^*(i) \qquad (11)$$

where $sim(q_a, d'_i)$ is the similarity between $q_a$ and $d'_i$ given by Eq. 3, $v^*(i)$ is the ranking of $r_i$ in the corpus.

We compute the matching scores (Eq. 11) between $q_a$ and all resources. We sort these scores in descending order and select top-K corresponding resources for recommendations.

Pseudo codes of the hybrid algorithm is described in Algorithm 1. In line 1, resource ranking is computed and stored in vector $v^*$. From line 2 to line 9, similarity between user profile $q_a$ and each resource $r_i$ (line 3-7) and their final matching score (line 8) are computed. After all, resources are sorted by their final matching scores (line 10) and top-K resources are picked up for recommendations (line 11).

**Algorithm 1.** Hybrid of content-based filtering and resource ranking

**input** : $q_a$: user profile of $u_a$, $R$: set of learning resources
**output**: $rec(a)$: recommendations for $u_a$

$v* \leftarrow$ PageRank $(R)$ ;
**foreach** $r_i \in R$ **do**
  $d_i \leftarrow$ Text description $(r_i)$ ;
  $d'_i \leftarrow$ Refinement of $d_i$ by $q_a$ by word semantic similarity ;
  $\overrightarrow{q_a} \leftarrow$ TF-IDF vector of $q_a$ ;
  $\overrightarrow{d'_i} \leftarrow$ TF-IDF vector of $d'_i$ ;
  $sim(q_a, d'_i) \leftarrow$ cosine $(\overrightarrow{q_a}, \overrightarrow{d'_i})$ ;
  $scr(q_a, r_i) \leftarrow sim(q_a, d'_i) \times v^*(i)$ ;
**end**
Sort $r_i \in R$ by $scr(q_a, r_i)$ in descending order. ;
$rec(a) \leftarrow$ top-K resources in the sorted list. ;

## 3 Experiment

We performed experiments on the learning resources that are published by the Open University of Humanities[6] (`http://www.uoh.fr/front`). However, due to the university privacy, historical usage data is not shared. So, we could not evaluate our approach based on ground-trust based metrics such as Precision/Recall, MAE, RMSE, etc. Instead, we measured the *similarity values*, the *coverage* of recommendations, the *convergence* of ranking vector and the *computation time* of our proposed algorithms in order to evaluate the *feasibility* of our approach. We elaborate in the following the collected dataset (section 3.1), our implementation (section 3.2) and experimental results (section 3.3).

### 3.1 Dataset

The Open University of Humanities is a French numerical university that provides open access to learning resources related to human science. These resources are created by teachers, lecturers of French higher educational schools. Each resource is published together with its description under the Learning Object Metadata (LOM) format. This description provides basic information of the resource such as title, abstract, keywords, discipline, types, creator, relations to other resources, etc.

As resource descriptions are public under a standard format, we crawled and parsed them to extract necessary information for our experiments. We collected 1294 resource descriptions, which indicate 62 publishers (universities, engineering schools, etc.), 14 pedagogic types (slide, animation, lecture, tutorial, etc.), 12 different formats (text/html, video/mpeg, application/pdf, etc), 10 different levels

[6] In French: Université ouverte des Humanités.

(school, secondary education, training, bac+1, bac+2, etc.) and 2 classification types (dewey, rameau). Among 1294 resources, 880 resources have relations with other resources, in which 692 resources have relation 'is part of', 333 resources have relation 'contains' and 573 resources have relation 'associates to'.

The collected dataset contains essential information for our proposed algorithms, including resource descriptions, which are used in the content-based filtering and word semantic similarity algorithms, and their relations, which are used in the resource ranking algorithm.

### 3.2 Implementation

We developed a Java program to crawl and extract the public resources. We used Apache Lucene[7] for word stemming and stop words removal. We used DISCO library[8] for word semantic similarity. We simulated 1000 queries, each of which includes keywords of 10 resources. We assumed that these resources are recently viewed by an active user. For each query, we computed recommendations in 6 different cases:

1. *SingularKW*: we consider keywords and resource descriptions as sets of singular keywords and run the content-based filtering algorithm.
2. *CompoundKW*: we preprocessed resource descriptions by identify their compound words that are matched with compound words in the query. Then, we consider each compound word as a singular word and run the content-based filtering algorithm.
3. *SemanticCB*: we replace words in resource descriptions by the most semantically similar words in the query. Then, we run the content-based filtering algorithm with the new resource descriptions.
4. *SingularKW-PageRank*: we run the hybrid algorithm that combines the content-based filtering with singular word matching and resource rankings based on their relations.
5. *CompoundKW-PageRank*: we run the hybrid algorithm that combines the content-based filtering with compound word matching and resource rankings.
6. *SemanticCB-PageRank*: we run the hybrid algorithm that combines the content-based filtering with semantic word similarity and resource ranking.

We run our proposed algorithms on a Mac laptop with the configuration as follows: CPU 2GHz Core i7, Memory 8G 1600 MHz, SSD 251G and OS X 10.9.2.

### 3.3 Results

We set damping factor $d = 0.85$ (like Google) and differential parameter $\alpha = 0.9$ (see section 2.3). We run our algorithms on 1000 different queries and obtained results as follows.

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

[8] `http://www.linguatools.de/disco/disco_en.html`

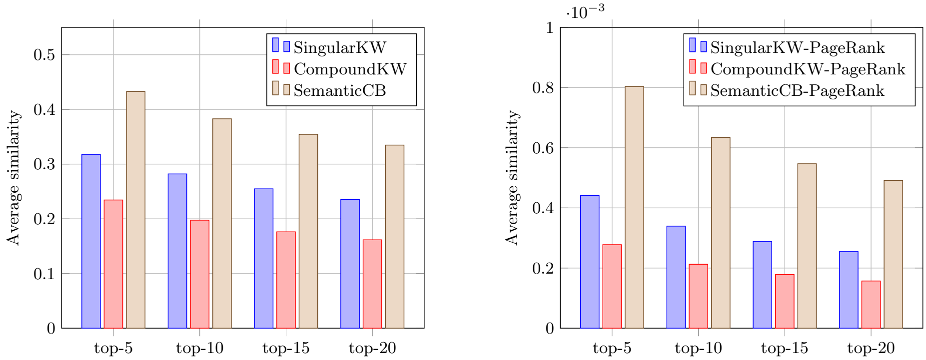



**Fig. 3.** Average similarity values with top-K selections

In the first experiment, we target to measure *similarity values* between queries and resources. For each query, we compute the average similarity values of top-K resources that are selected for recommendations. Fig. 3 shows experimental results of the 6 different cases mentioned above. The SemanticCB and SemanticCB-PageRank cases achieve the highest average similarity values as they take into account both syntactic and semantic word similarity. Meanwhile, The CompoundKW and CompoundKW-PageRank have the lowest similarity values as they consider only syntactical matching between compound words which leads to the smallest number of matching pairs. The cases with PageRank algorithm have very small similarity values because the ranking values of resources are very small to satisfy that sum of all of them is equal to 1 (Max.=$7.16 \times 10^{-3}$, Min.=$1.16 \times 10^{-4}$).

In the second experiment, we measure the *coverage* of recommendations, i.e. the percentage of resources that are considered for the top-K selection. We obtained that the CompoundKW and CompoundKW-PageRank cases (notated by CompoundKW*) have the lowest coverage (Fig. 4). It is because the number of compound word matchings is much smaller than the number of singular word matchings and word semantic matchings. We also obtained that the coverage of the SemanticCB* cases is always 1. It means that for each query, we always find at least a word in a resource description that is semantically matched to a word in the query with a matching value greater than 0.

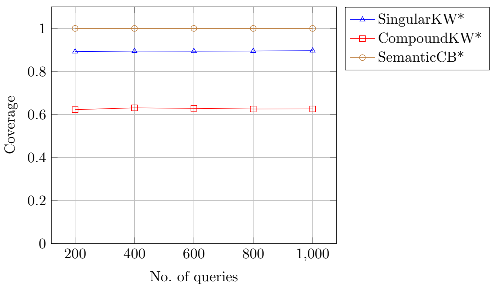



**Fig. 4.** Coverage of recommendations

In the third experiment, we target to measure the *convergence* of resource rankings with different thresholds (from $10^{-4}$ to $10^{-9}$). Fig. 5 shows the number

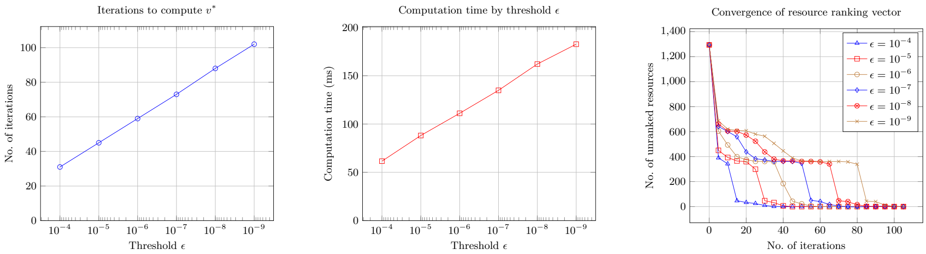



**Fig. 5.** Experiments on resource ranking

of iterations needs to be performed to compute the ranking vector $v^*$, the corresponding computation times and the convergence of resource rankings. Theses results show that our approach can rapidly rank resources based on their relations, for instance, we can rank 1294 resources within 180ms with a very small threshold $10^{-9}$.

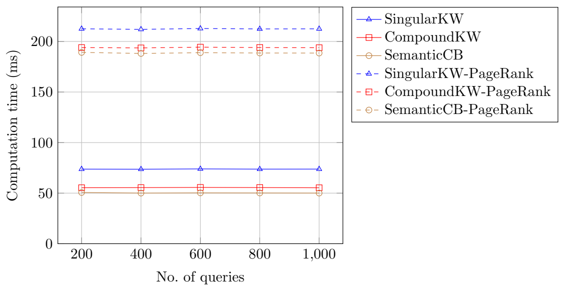



**Fig. 6.** Computation time with $\epsilon = 10^{-6}$

In the last experiment, we target to measure the *computation time* of our algorithms, without the data preprocessing time. Fig. 6 shows that the SemanticCB* cases have the smallest computation time while the SingularKW* cases have the highest computation time. It is because the number of dimensions in the resource vector space in the SemanticCB* cases is the smallest and in the SingularKW* cases is the highest. The cases with PageRank have much more higher computation time than others as they include the computation time of resource ranking. Fig. 6 also shows that our algorithms can make recommendations in very short time (around 200 ms with 1294 resources).

According to the limitation of dataset, we do not evaluate our approach using ground-trust based metrics. However, experimental results on the *similarity values*, the *coverage* of recommendations, the *convergence* of ranking vector and the *computation time* showed that our approach is able to make recommendations in a *very short time* and *for all queries*. This means that our approach is *feasible* in reality.

## 4 Related Work

On the TEL domain, a number of recommendation approaches have been proposed to encourage the usage of learning resources for online education. Manouselis et. al. [14] have made a deep survey on these existing approaches, which apply recommendation techniques on different online education contexts.

Common used techniques such as collaborative filtering [12,20], content-based filtering [9,11], association rules [12,19], user ratings [5,15] and feedback [8] analysis have been exploited. However, none of existing approaches considers a combination of syntactic and semantic matching. In addition, most of them still remain at a design or prototyping stage. Only few of them have been reported to be evaluated through trials that involved human users [14]. In our approach, we take into account both syntactic and semantic matching together with resource ranking. We also provide experiments on a dataset of real online learning resources.

A related work that applied content-based and collaborative filtering on recent viewed resources has been proposed by Khribi et. al. [9]. They also presented experiments on resources that are presented in the standard Learning Object Metadata (LOM) format. However, different from them, we present another combination of existing recommendation techniques and we consider historical keywords as user profile instead of entire resource content. Although we performed experiments on LOM formatted resources, our approach can be applied on different resource formats as we take in to account resource descriptions instead of their formats.

Another related work that considered word similarity has been proposed by Chen et. al. [3]. They apply word similarity to compute the matching between user query and web contents. However, in their approach, they proposed to expand the user query by including all of their similar words. In our approach, instead of expanding the user query, we replacing words in a resource description by their most semantically similar words in the query.

## 5 Conclusion

Recommender systems have been considered as a pivotal solution to encourage the usage of online learning resources. In this paper, we present an innovative hybrid approach that combines three recommendation techniques: collaborative-filtering, semantic similarity and page rankings to generate resource recommendations. By this combination, our approach is able to recommend important resources that are syntactically and semantically relevant to user requests. In our approach, user profile is implicitly built based on keywords of recent viewed resources. Hence, we do not ask any effort from users. In addition, as recent viewed resources present recent interest of user, our approach is able to recommend resources that are close to user interest.

In future work, we will validate our approach on other datasets using ground-truth based metrics such as precision/recall, MAE and RMSE. We will take into account resource levels in order to filter resources that are best fitted to the learner's level. We also plan to integrate collaborative filtering and clustering techniques to improve the quality of recommendations.

**Acknowledgments.** This work has been fully supported by the French General Commission for Investment (Commissariat Général à l'Investissement), the Deposits and Consignments Fund (Caisse des Dépôts et Consignations) and the Ministry of Higher Education & Research (Ministère de l'Enseignement Supérieur et de la Recherche) within the context of the PERICLES project (`http://www.e-pericles.org`).

## References

1. Avancini, H., Straccia, U.: User recommendation for collaborative and personalised digital archives. Int. J. Web Based Communities 1(2), 163–175 (2005)
2. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. In: WWW7, pp. 107–117. Elsevier Science Publishers B. V. (1998)
3. Chen, L., Sycara, K.: Webmate: A personal agent for browsing and searching. In: AGENTS 1998, pp. 132–139. ACM, New York (1998)
4. Danushka, B., Yutaka, M., Mitsuru, I.: Measuring semantic similarity between words using web search engines. In: WWW 2007, pp. 757–766. ACM (2007)
5. Drachsler, H., Pecceu, D., Arts, T., Hutten, E., Rutledge, L., van Rosmalen, P., Hummel, H., Koper, R.: Remashed — recommendations for mash-up personal learning environments. In: Cress, U., Dimitrova, V., Specht, M. (eds.) EC-TEL 2009. LNCS, vol. 5794, pp. 788–793. Springer, Heidelberg (2009)
6. Huang, Y.M., Huang, T.C., Wang, K.T., Hwang, W.Y.: A markov-based recommendation model for exploring the transfer of learning on the web. Educational Technology & Society 12(2), 144–162 (2009)
7. Hummel, H.G.K., van den Berg, B., Berlanga, A.J., Drachsler, H., Janssen, J., Nadolski, R., Koper, R.: Combining social-based and information-based approaches for personalised recommendation on sequencing learning activities. IJLT (2007)
8. Janssen, J., Tattersall, C., Waterink, W., van den Berg, B., van Es, R., Bolman, C., Koper, R.: Self-organising navigational support in lifelong learning: How predecessors can lead the way. Comput. Educ. 49(3), 781–793 (2007)
9. Khribi, M.K., Jemni, M., Nasraoui, O.: Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval. Educational Technology & Society 12(4), 30–42 (2009)
10. Kolb, P.: Experiments on the difference between semantic similarity and relatedness. In: Jokinen, K., Bick, E. (eds.) NODALIDA 2009, vol. 4, pp. 81–88 (2009)
11. Koutrika, G., Ikeda, R., Bercovitz, B., Garcia-Molina, H.: Flexible recommendations over rich data. In: RecSys 2008, pp. 203–210. ACM (2008)
12. Lemire, D., Boley, H., McGrath, S., Ball, M.: Collaborative filtering and inference rules for context-aware learning object recommendation. International Journal of Interactive Technology and Smart Education 2(3) (August 2005)
13. Li, Y., Bandar, Z.A., McLean, D.: An approach for measuring semantic similarity between words using multiple information sources. IEEE Trans. on Knowl. and Data Eng. 15(4), 871–882 (2003)
14. 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, pp. 387–415 (2011)
15. Manouselis, N., Vuorikari, R., Assche, F.V.: Simulated analysis of maut collaborative filtering for learning object recommendation. In: SIRTEL 2007 (2007)
16. Nadolski, R.J., van den Berg, B., Berlanga, A.J., Drachsler, H., Hummel, H.G., Koper, R., Sloep, P.B.: Simulating light-weight personalised recommender systems in learning networks: A case for pedagogy-oriented and rating-based hybrid recommendation strategies. J. of Artificial Societies and Social Simulation (2009)

17. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab (1999)
18. Richardson, R., Smeaton, A., Murphy, J.: Using wordnet as a knowledge base for measuring semantic similarity between words. In: AICS 2009 (1994)
19. Shen, L.-p., Shen, R.-M.: Learning content recommendation service based-on simple sequencing specification. In: Liu, W., Shi, Y., Li, Q. (eds.) ICWL 2004. LNCS, vol. 3143, pp. 363–370. Springer, Heidelberg (2004)
20. Tang, T., McCalla, G.: Smart recommendation for an evolving e-learning system: Architecture and experiment. I. J. on E-Learning, 105–129 (2005)
21. Tzikopoulos, A., Manouselis, N., Vuorikari, R.: An overview of learning object repositories. In: Erickson, J. (ed.) Database Technologies: Concepts, Methodologies, Tools, and Applications, pp. 362–383. IGI Global (2009)
22. Wills, R.S.: Google's pagerank: The math behind the search engine. Math. Intelligencer, 6–10 (2006)

# Usage-Based Clustering of Learning Resources to Improve Recommendations

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 usage-based technique for clustering learning resources accessed in online learning portals. This approach solely relies on the usage of the learning resources and does not consider their content or the relations between the users and the resources. In order to cluster the resources, we calculate higher-order co-occurrences, a technique taken from corpus-driven lexicology where it is used to cluster words based on their usage in language. We first outline how we adapt the approach to then present an extensive evaluation that shows the effects of the clustering. Finally, we show how the resulting clusters can be used to enhance recommender systems.

**Keywords:** clustering, higher-order co-occurrences, usage data, recommender systems, educational data mining, learning analytics.

## 1 Introduction

Data sets coming from the domain of technology enhanced learning (TEL) often suffer from sparsity, e.g. in respect to semantic metadata describing the learning resources or available rating data. This makes it difficult to find relations between the resources and to create recommendations for learners and teachers. After the analysis of the data sets submitted to the first dataTEL[1] Challenge, Verbert et al. [1] conclude that further research on similarity measures and implicit relevance indicators is required to sufficiently support the learners and teachers.

Commonly, similarity measures are based on the resources' content representations which hold semantic features that are manually or automatically created. Automatic extraction techniques analyse the textual content of a learning resource, e.g. to generate a weighted list of characteristic keywords. Though, many repositories do not only hold textual learning resources but also images, videos, or audio files for which the automatic extraction of semantic features is significantly more difficult or even impossible. Furthermore, many web portals offer learning resources in different languages which makes it complex to compare the keywords and requires additional language modules. An alternative to the automatic extraction of semantic features is their manual creation by experts.

[1] http://www.teleurope.eu/pg/groups/9405/datatel/

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 317–330, 2014.
© Springer International Publishing Switzerland 2014

While this approach is common in traditional environments like libraries and produces high quality metadata, it is also expensive in terms of time and money. Further, maintenance of the data is problematic as well. Nowadays, the use of user-generated, i.e. social metadata like tags and comments is popular. Tags in particular provide an effective way to represent user interests and to help users in finding resources about a specific topic independent of their media type [6, 7]. However, in the given setting, tags holding the same meaning might be provided in different languages which makes it difficult to compare them. Furthermore, tags can be unique, i.e. only used once and it can be assumed that a reasonable number of resources might not hold any tags at all.

In this paper, we present a technique of detecting similarity relations between learning resources by clustering them solely based on their usage. For this purpose, we apply the technique of calculating higher-order co-occurrences to learning resources used in web portals. This way, only the learners' interactions with the resources are taken into account but no further information about the resources' contents are required.

The paper is structured as follows. Chapter 2 provides an overview on the background of higher-order co-occurrence clustering that has tradition in corpus linguistics and shows how this approach can be applied to learning resources and their usage. Chapter 3 presents the experimental set-up of the evaluation and chapter 4 provides an extensive discussion of the evaluation results. Finally, chapter 5 gives a conclusion and an outlook on future work.

## 2 Higher-order Co-occurrence Clustering

### 2.1 Linguistic Background

The calculation of higher-order co-occurrences is used in corpus linguistics to cluster lexical entities (e.g. words) into semantically similar groups based on their usage in language. The first-order co-occurrences of a lexical entity can be calculated by taking the context of that entity into account, e.g. the sentences in which it occurs. Here, the significant co-occurrences of an entity form its first-order co-occurrence class and entities which co-occur in first-order co-occurrence classes are second-order co-occurrences. These second-order co-occurrences again can be used as input to calculate third-order co-occurrences and so forth.

When this procedure is repeated several times, entities that do not hold a sufficient number of significant relations to other entities are excluded from a certain order on. However, the remaining higher-order co-occurrence classes tend to get stable, i.e. their elements do not change any more. This indicates that there exist universal relations between the entities in the remaining classes that induce their aggregation again in each iteration step. In fact, these stable higher-order co-occurrence classes have shown to usually hold semantically related entities [2]

For example, the first-order co-occurrence class of the company name *IBM* that is created by analysing the text corpora offered by the portal wortschatz.uni-leipzig.de is rather heterogeneous. It contains lexical entities like *computer manufacturer*, *stock exchange*, *global* etc. Though, after some iterations the higher-order co-occurrence classes become more stable and homogeneous. Finally, the co-occurrence class of tenth order only contains names of computer-related companies like *Microsoft*, *Sony* etc. [2]. These classes can then for example be used as input for the extension of ontologies or the categorisation of texts [3][4].

## 2.2 Transfer of the Approach to Learning Resources

While lexical entities are used in sentences, learning resources accessed in web portals are used in user sessions. A user session comprises all resources accessed by a user between a log-in and log-out event or in a pre-defined time frame. Thus, on a more abstract level, in both cases, entities are used in sequences of different lengths. Therefore, we can transfer the technique of higher-order co-occurrence clustering to learning resources if the needed usage data is given. In order to clarify the clustering process with an example, the following user sessions $US_1$, $US_2$, and $US_3$ serve as input for the creation of higher-order co-occurrences.

$US_1 = \langle A, B, C, D \rangle$
$US_2 = \langle A, B, D \rangle$
$US_3 = \langle D, E, F \rangle$

The calculation of the first-order co-occurrences results in the following (temporary) classes for $A$, $B$, $C$, $D$, $E$, and $F$. The total frequencies of the resources and their co-occurrences are given in brackets, e.g. $A$ and $D$ co-occur two times whereas $A$ is used in two user sessions and $D$ is used in three user sessions.

$A(2) \rightarrow B(2), C(1), D(2)$
$B(2) \rightarrow A(2), C(1), D(2)$
$C(1) \rightarrow A(1), B(1), D(1)$
$D(3) \rightarrow A(2), B(2), C(1), E(1), F(1)$
$E(1) \rightarrow D(1), F(1)$
$F(1) \rightarrow D(1), E(1)$

The given frequencies are used to calculate the significance values for each resource's co-occurrences. For demonstration purposes, exemplary significance values are used in this example. (For successfully applying measures of significance as presented in section 3.2, a larger collection of user sessions is required.) The exemplary significance values of the co-occurrences are given in brackets.

$A \rightarrow B(1.7), C(1.4), D(1.7)$
$B \rightarrow A(1.7), C(1.4), D(1.7)$
$C \rightarrow A(1.4), B(1.4), D(1.1)$
$D \rightarrow A(1.7), B(1.7), C(1.1), E(1.1), F(1.1)$
$E \rightarrow D(1.1), F(1.4)$
$F \rightarrow D(1.1), E(1.4)$

In order to identify the insignificant co-occurrences resource-specific thresholds must be calculated. For simplification, a general threshold of 1.3 is applied in this example. Thus, the final co-occurrence classes can be generated by the deletion of all co-occurrences with a significance value lower than 1.3.

First-order co-occurrence class for $A$: $\{B, C, D\}$
First-order co-occurrence class for $B$: $\{A, C, D\}$
First-order co-occurrence class for $C$: $\{A, B\}$
First-order co-occurrence class for $D$: $\{A, B\}$
First-order co-occurrence class for $E$: $\{F\}$
First-order co-occurrence class for $F$: $\{E\}$

Using these first-order co-occurrence classes as input, the following temporary second-order co-occurrences classes arise for $A$, $B$, $C$, and $D$. The second-order co-occurrence classes for $E$ and $F$ are empty. This means they do not have significant second-order co-occurrences and, thus, are excluded from the further iteration steps. The total frequencies of the co-occurrences and the associated exemplary significance values are given in brackets.

$A(3) \rightarrow B(2; 1.5), C(1; 1.3), D(1; 1.3)$
$B(3) \rightarrow A(2; 1.5), C(1; 1.3), D(1; 1.3)$
$C(2) \rightarrow A(1; 1.3), B(1; 1.3), D(2; 1.7)$
$D(2) \rightarrow A(1; 1.3), B(1; 1.3), C(2; 1.7)$

Using again a threshold of 1.3, the following second-order co-occurrence classes arise.

Second-order co-occurrence class for $A$: $\{B, C, D\}$
Second-order co-occurrence class for $B$: $\{A, C, D\}$
Second-order co-occurrence class for $C$: $\{A, B, D\}$
Second-order co-occurrence class for $D$: $\{A, C, D\}$

These classes can now be used as input for the calculation of third-order co-occurrence classes and so forth. The calculation stops when the classes get stable, i.e. they do not change any more in further iterations. These stable higher-order co-occurrence classes are also called *clusters* in the remainder of this paper. It must be noted that this procedure generates a fuzzy clustering without membership values, i.e. a resource can fully belong to several clusters.

## 3 Experimental Set-up

### 3.1 Data Set: MACE

The MACE[2] (Metadata for Architectural Contents in Europe) project relates digital learning resources about architecture with each other across repository

[2] `http://mace-project.eu/`

boundaries [5]. While interacting with the MACE portal, users are monitored and their activities are recorded as CAM (Contextualised Attention Metadata, [6]) instances. The considered event types comprise accessing the metadata of a learning resource in the MACE portal and accessing the learning resource in its origin repository as well as metadata provision activities, i.e. rating, tagging, and commenting. 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 resource. The CAM instances for the MACE data set were collected in a period of three years from September 2009 to October 2012. The data set holds 117,907 events on 12,442 learning resources conducted by 630 registered users. Here, a user session is defined to hold all activities conducted by a user without a break of more than an hour. This way, the events are split in 4,219 sessions in which each resource is used in on average 2,17 sessions.

Furthermore, 78.69% of the learning resources hold additional semantic metadata, i.e. 70.8% hold free-text tags added by users, 14.83% hold classifications from a controlled vocabulary added by experts, and 8.82% hold both. Each tagged resource holds on average 6.59 tags and each classified resource holds on average 2.27 classifications. This information is used to calculate semantic metadata-based similarities between the learning resources that serve as baseline for the evaluation of the clusterings. Because each tag and classification can only be assigned once to a resource in the MACE system, i.e. no weighting is conducted, a set-based approach is applied and the Jaccard coefficient is calculated to measure the semantic metadata-based similarity of two resources [2].

Additionally, 429 learning resources were rated by 76 users, each of these learning resources was rated at least once and at maximum 4 times (on average 1.11 times), and each of the 76 users rated 1-91 learning resources (on average 6.52). This results in a user-item-rating matrix with a sparsity of 98.54%. This rating data is used to evaluate the ability of the clustering to improve existing recommendation approaches.

### 3.2 Significant Co-occurrences

We define two resources to be co-occurrences if they co-occur in at least one user session and use statistical association measures to calculate their significance. Basic association measures calculate a significance score by comparing the observed frequency of a co-occurrence with its expected frequency, e.g. Mutual Information (MI) or z-scores [7]. More sophisticated association measures and independence tests are always based on a cross-classification of a set of resources, e.g. using contingency tables. Examples are log likelihood (LL) and corrected $\chi^2$ (cor-$\chi^2$) as well as the Poisson-based similarity (PAM) [7]. In contrast to the more simple approaches, they do not only consider the expected co-occurrence frequency of the two resources but compute the expected frequencies for all cells in the contingency table. In this paper, the association measures MI, cor-$\chi^2$, LL, and PAM are used to calculate the significance of co-occurrences.

After the calculation of the co-occurrences' significance values, the most significant ones must be selected for each resource. Since there is no standard scale of measurement to draw a clear distinction between significant and non-significant occurrences, there are two ways to do so, i.e. by selecting only the $n$ most significant co-occurrences for each resource or by using a threshold. Here, selecting the best $n$ would imply to have a pre-defined cluster size which is not desirable, thus, a threshold is used. Because the calculated significance scores for resource pairs are only comparable if they have one resource in common, a resource-specific threshold, i.e. its average significance, is used to distinguish between relevant and non-relevant co-occurrences. If the results show that too many co-occurrences are neglected leading to missing relations between the resources and thus to resources not being clustered at all, the resource-specific thresholds can be multiplied with a constant $\alpha$ holding any value between 0 and 1. However, the $\alpha$ value should be as high as possible to not blur the relations between the items which would lead to large and fuzzy clusters. Using the MACE data set, more than 90% of the resources cannot be clustered but are filtered using the mean values, thus, they are multiplied with an $\alpha$ value of 0.95 which forces less than 10% of the resources to be filtered.

### 3.3 Cluster-Based Recommendations

There are several applications in which the clusterings can be used to support students and teachers, a striking one is their usage in recommender systems. The most straightforward approach is to display resources that are similar to the one that is currently accessed, i.e. resources from the same cluster. Additionally, the clusters can be used to predict ratings based on the available user rating behaviour which can be explicit or implicit. Equation 1 shows how a predicted rating $\hat{r}(u,i)$ of user $u$ on resource $i$ can be calculated with $P(u)$ being the profile of user $u$ holding all ratings on other resources. The similarity $sim(i,j)$ of the two resources $i$ and $j$ is 1 if they belong to the same cluster and 0 otherwise.

$$\hat{r}(u,i) = \frac{\sum_{j \in P(u), i \neq j} (sim(i,j) * r(u,j))}{\sum_{j \in P(u), i \neq j} |sim(i,j)|} \tag{1}$$

Additionally to the cluster-based recommendation approach, the predicted ratings can be used to boost existing collaborative filtering techniques. The cluster-boosted collaborative filtering is a two-step procedure, first, the given user-item-rating matrix, which is usually very sparse, is filled with ratings predicted based on the resources' cluster memberships. Second, the filled matrix is used by standard collaborative filtering approaches that benefit from the additional data.

### 3.4 Evaluation Process

The evaluation of clustering results can be intrinsic or extrinsic. An intrinsic evaluation metric measures how close the resources within the clusters are among each other in relation to how distant they are to resources in other clusters. This means that in order to conduct an intrinsic evaluation, pair-wise similarities of the resources must be given. An extrinsic evaluation can be conducted by using a clustering in a concrete application and measuring its additional benefit or by comparing it against a *gold standard*, i.e. a given clustering that is created by domain experts [8]. Here, no *gold standard* is given but pair-wise resource similarities that are based on the semantic metadata are available. Thus, the clusterings are first evaluated using the intrinsic evaluation metrics presented in the next paragraphs. The best clustering is then used to enhance recommender systems in order to conduct an extrinsic evaluation in a concrete setting.

**Intrinsic Evaluation.** The Davies-Bouldin (DB) index was originally developed to determine the proper number of clusters in the given data by relating the intra-cluster distances (which should be as low as possible) and the inter-cluster distances (that should be as high as possible) in order to find a suitable trade-off [9]. Equation 2 shows the calculation of the Davies-Bouldin index with $\sigma(C_i)$ and $\sigma(C_j)$ being the dispersions of clusters $C_i$ and $C_j$, respectively, and $d(C_i, C_j)$ being the distance between the two clusters. The dispersion $\sigma$ of a cluster is calculated as the average distance of the clusters' resources to its centroid while the distance between two clusters is calculated as the distance of their centroids using the resources' semantic metadata (see section 3.1). The better the clustering based on these criteria, the smaller the Davies-Bouldin index. However, the Davies-Bouldin index can only be used to compare clusterings on the same data set and cannot be interpreted as standalone value.

$$\mathrm{DB} = \frac{1}{n} \sum_{i=1}^{n} \max_{1 \leq j \leq n, i \neq j} \left( \frac{\sigma(C_i) + \sigma(C_j)}{d(C_i, C_j)} \right) \tag{2}$$

An evaluation metric that tries to determine whether resources are correctly positioned or classified without the need of a gold standard to compare the clustering to is proposed in [10]. A resource $i$ is assumed to be correctly positioned if its intra-cluster similarity, i.e. the accumulated similarities to the resources in its cluster (see equation 3), is higher than its inter-cluster similarity, i.e. the accumulated similarities to all resources that do not belong to this cluster (see equation 4). Here, the value for this metric is calculated as the percentage of correctly positioned items (see equation 5).

$$\text{intra-sim}(i) = \sum_{j \in C} sim(i, j), \text{ with } i \in C \tag{3}$$

$$\text{inter-sim}(i) = \sum_{j \notin C} sim(i, j), \text{ with } i \in C \tag{4}$$

$$\mathrm{CPI} = \frac{100 * \sum_{i \in I} \mathrm{CP}(i)}{|I|}, \text{ with } \mathrm{CP}(i) = \begin{cases} 1, \text{intra-sim}(i) > \text{inter-sim}(i) \\ 0, \text{else} \end{cases} \tag{5}$$

Vries et al. [11] propose to use the divergence from a random baseline to filter ineffective clusterings, i.e. those clusterings that achieve high scores using any evaluation metric but provide no value as clustering solutions. In opposite, they define a useful clustering to work beyond a random approach. It is important that this randomly created clustering holds the same cluster size distribution as the one it is compared to because some evaluation metrics have a bias towards many small or few large clusters. For example, a clustering that puts each resource in a single cluster will receive the highest possible score for purity, as no cluster holds resources that belong to different topics. Though, it cannot be regarded as a useful clustering and, thus, it will not perform better compared to a random baseline holding the same cluster size distribution.

**Extrinsic Evaluation.** In order to evaluate whether the clusterings provide a real value in an application that supports students, they are used to create recommendations and to enhance existing recommender systems as discussed in section 3.3. The performance of the recommendation techniques is measured in terms of Root Mean Squared Error (RMSE), see formula 6, which is calculated in a 5-fold cross evaluation. The RMSE measures the deviation between a predicted rating $\hat{r}(u,i)$ and the user's true rating $r(u,i)$ for all users $U$ and their ratings $P(u)$ in the test set. Furthermore, the RMSE squares the deviations before they are averaged, thus giving a higher weight to large errors. Additionally to the RMSE, the coverage, i.e. the number of user-item pairs in the test set for which a rating can be predicted, is calculated

$$\text{RMSE} = \sqrt{\frac{\sum_{u \in U} \sum_{i \in P(u)} (p(u,i) - r(u,i))^2}{\sum_{u \in U} |P(u)|}} \tag{6}$$

# 4 Results

## 4.1 Clustering Process

Figure 1 shows the evolution of the mean cluster sizes on the MACE data set. In the first iteration, the cluster sizes are dependent of the chosen association measure and vary from 67 (cor-$\chi^2$) to 168 (MI). In the next iterations the clusters grow and reach their maximum size between the fourth and eighth iteration. Thereafter, the clusters start to shrink again until they reach a stable size.

This behaviour can be motivated by the fact that many resource pairs are distantly connected, thus, the cluster sizes increase at the beginning. However, only pairs that are connected by a significant number of paths of different length are aggregated again in each iteration step, thus, the cluster sizes start to decrease again from a certain order on until they get stable. When the clustering is based on MI or LL, the mean cluster sizes increase more in the iteration process than for the $\chi^2$- or PAM-based clusterings. This is due to the clusterings' different cluster size distributions. For example, when MI is used, many small (less than 10 resources) and few very large (more than 1,000 resources) clusters are produced. This means until the final large clusters are formed, the clusters' constituents, i.e. many large (and overlapping) clusters, are part of the clustering

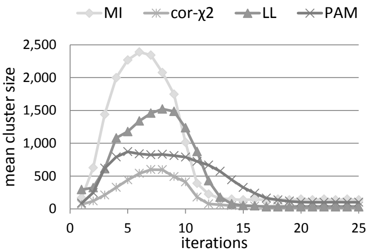



**Fig. 1.** Mean Cluster Size

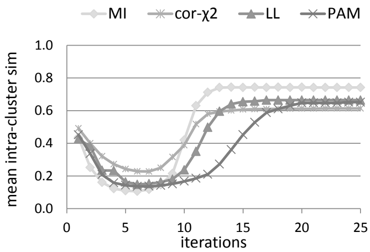



**Fig. 2.** Mean Intra-Cluster Similarity

as well and increase the average cluster size and, thus, the memory need and processing time. The final mean cluster sizes vary from 30 (LL) to 137 (MI).

Figure 2 shows the mean intra-cluster similarities using the different association measures. The intra-cluster similarity of a cluster is calculated as the average similarity of its resources to its centroid. The intra-cluster similarities of the first-order classes are very similar with a value of approximately 0.44. In the next iteration steps, the mean cluster sizes increase and force a drop of the corresponding intra-cluster similarities. Thereafter, when the cluster sizes start to decrease again, the intra-cluster similarities increase accordingly until they get stable. For MACE, MI is the best performing association metric with a mean intra-cluster similarity of 0.7416 followed by LL (0.6643), PAM (0.6518), and cor-$\chi^2$ (0.6133).

## 4.2 Cluster Merging

The basic assumption is that a resource can fully belong to several clusters. However, the analyses of the clusters show that the fuzziest clusterings are also the ones with the lowest intra-cluster similarity and vice versa. Thus, this section investigates if the quality of a clustering can be increased by merging clusters that overlap by a certain degree. The quality is again measured by the mean intra-cluster similarity since the Davies-Bouldin index rewards well-isolated clusters and the percentage of correctly positioned items has a bias against large clusters. This means that the resulting values of these measures can be assumed to improve as soon as overlapping clusters are merged. In opposite, the mean intra-cluster similarity has a bias against small cluster, thus, if the intra-cluster similarity increases when merging the clusters, it can be seen as an indication that the clustering's quality is improved.

Figure 3 shows the intra-cluster similarities for varying Jaccard coefficients that cause two clusters to be merged. The Jaccard coefficient is calculated as the intersection size of two clusters divided by their union size. In fact, the lower the Jaccard coefficient, the higher the intra-cluster similarity, i.e. the best results are reached if two clusters are merged as soon as they have one overlapping resource. It can be observed that in most cases all overlapping clusters are merged into one single large cluster. Since the intra-cluster similarity raises through this process

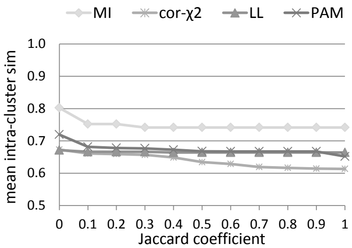



**Fig. 3.** Mean Intra-Cluster Similarity after Merging

it can be assumed that the merging affects primarily those clusters with low intra-cluster similarities and by merging them - which typically results in one large cluster with a low intra-cluster similarity - they are only included once in the calculation of the mean intra-cluster similarity. However, according to [8] it is preferable to only have one "rag bag", i.e. a cluster holding those resources that cannot be clustered more reasonably. Thus, the merged clusters (with a Jaccard index $> 0$) are evaluated in the next sections.

### 4.3 Intrinsic Evaluation

Table 1 shows the intra-cluster similarity (Sim), the Davies-Bouldin (DB) index, and the percentage of correctly positioned items (CPI) for the four final clusterings. The intra-cluster similarity can take any value between 0 and 1 (with 1 being best), the DB index can take any positive value (the smaller the value the better the clustering), and the CPI can take any value between 0 and 100 (with 100 being the best). In order to calculate the baselines, for each clustering ten associated baseline clusterings are randomly created while considering the respective cluster size distribution. The evaluation results for these ten baseline clusterings are averaged and the standard deviation is given in the table as well. It can be noted that all four clusterings score better than their associated baselines in respect to the three evaluation metrics.

**Mean Intra-cluster Similarity.** The usage of MI produces the clustering with the highest intra-cluster similarity. This can partly be explained by the cluster size distribution, i.e. many very small clusters which by definition have a high intra-cluster similarity and few very large clusters that carry not much weight. This also explains why MI has the highest baseline. However, the intra-cluster similarity for MI exceeds its baseline by 56.6% which indicates that the clustering provides a real value. The other measures also exceed their baseline by at least 35.41% (LL) with PAM being the second and cor-$\chi^2$ being the third best measure.

**Davies-Bouldin Index.** The DB index assesses how dense and well-separated the clusters are. Similarly to the intra-cluster similarity, MI and PAM perform

**Table 1.** Evaluation results with baselines (MACE)

| Measure | Sim | Sim Baseline | DB | DB Baseline | CPI | CPI Baseline |
|---|---|---|---|---|---|---|
| MI | 0.8023 | 0.5123 $\sigma = 0.0079$ | 1.3363 | 6.5790 $\sigma = 0.5610$ | 65.78% | 56.15% $\sigma = 0.12\%$ |
| cor-$\chi^2$ | 0.6731 | 0.4486 $\sigma = 0.0023$ | 2.3373 | 4.6786 $\sigma = 0.0702$ | 14.23% | 0.31% $\sigma = 0.05\%$ |
| LL | 0.6714 | 0.4958 $\sigma = 0.0054$ | 1.7877 | 4.3675 $\sigma = 0.1067$ | 31.20% | 1.11% $\sigma = 0.12\%$ |
| PAM | 0.7194 | 0.4614 $\sigma = 0.0061$ | 1.9579 | 6.9065 $\sigma = 0.1255$ | 44.81% | 9.47% $\sigma = 0.90\%$ |

best when comparing the measures among each other and they also show the highest divergence to the baseline. However, LL and cor-$\chi^2$ also deviate from the baseline by at least half of their values.

**Correctly Positioned Items.** It can be noticed that the CPI baseline is far higher for MI than for the other association measures. This is because additionally to very small clusters, MI also creates one large cluster holding all resources that cannot be clustered otherwise and the likelihood that a resource contained in this large cluster is positioned correctly according to this measure is very high. However, the CPI value of the higher-order co-occurrence clustering using MI is still higher than for the baseline which shows that it has a positive effect on the clustering. The next best clustering is again the PAM-based one with the highest divergence to the baseline followed by the LL- and cor-$\chi^2$-based clusterings.

**Association Measures.** The results show that independent from the association measure used the higher-order co-occurrence clustering is able to group similar resources in well-separated clusters. However, the rag bags produced in the clustering process can grow very large. For example, the clustering produced using MI holds the best values in respect to all three evaluation metrics. Though, it contains one large cluster that holds 9,261 resources, i.e. 77.19% of all clustered resources. This means that even if the clustering is very good for a part of the resources, a large part of the resources can be considered to be not clustered at all. This phenomenon can be found for the other association measures as well even if not as distinctively as for MI. The largest clusters of the other clusterings reach a maximum size of 5,681 (LL), 3,417 (PAM), and 1,949 (cor-$\chi^2$). It is striking that the usage of cor-$\chi^2$ leads to smaller rag bags, which is an advantage, but also to worse evaluation results than the other association measures. However, in a real world setting, cor-$\chi^2$ might perform better than other measures since it creates relations between a large part of the resources.

**Table 2.** Cluster-based Recommendation

| | **RMSE** | **Coverage** |
|---|---|---|
| MI | 0.2009 | 20.94% |
| cor-$\chi^2$ | 0.7219 | 50.94% |
| LL | 0.8699 | 53.21% |
| PAM | 0.7767 | 46.79% |

The results of the intrinsic evaluation indicate that the clusterings provide a real value. Additionally, it can be seen that the choice of the association measure strongly influences the structure of the clusterings. When the clustering is based on the association measure MI, a large part of the resources is collected in a rag bag and not reasonably clustered while the remaining clustering shows the highest quality compared to the other ones. On the contrary, the quality of the clustering created using cor-$\chi^2$ is not as convincing but most resources are reasonably clustered which might be an advantage in a real setting.

### 4.4 Extrinsic Evaluation

Table 2 shows the RMSE and the coverage for the predicted ratings calculated using the different clusterings and the formula shown in section 3.3. The usage of the MI-based clustering results in the smallest error, i.e. 0.2009, however, it can only predict ratings for 20.94% of the user-item pairs in the test set. This result is in line with the findings of the intrinsic evaluation, i.e. using MI only a subset of the resources can be clustered but the clusters that can be created hold the best values in terms of mean intra-cluster similarity, DB index, and CPI. All other association measures create clusterings that can be used to predict about 50% of the ratings in which cor-$\chi^2$ holds the second-lowest RMSE with 0.7219, followed by PAM (0.7767) and LL (0.8699). Here, the usage of cor-$\chi^2$ can be considered to be the best trade-off between quality and quantity. Thus, in the following, the clustering created using the cor-$\chi^2$ is used to boost the existing collaborative filtering approaches.

Table 3 shows the results of the selected collaborative filtering (CF) approaches and their cluster-boosted equivalents. Here, the CF approaches are item-based CF (IBCF) and user-based CF (UBCF) as well as the Single Value Decomposition (SVD) from the MyMediaLite Recommender System Library [12] and the Biased Matrix Factorisation from the PREA toolkit [13]. It can be observed that the neighbourhood-based CF approaches, i.e. IBCF and UBCF, perform similarly to the MI-based clustering in terms of coverage but produce higher errors, i.e. 0.5776 (UBCF) and 1.0814 (IBCF). The SVD also suffers from the sparsity and predicts ratings for only 30% of the user-item pairs in the test set with an RMSE of 0.8303. The BMF is an exception as it always returns a predicted rating even if the user or the resource are not part of the training set which results in the highest RMSE, i.e. 1.1947.

**Table 3.** MACE: Cluster-based Filtering

| | Baseline | | Cluster-boosted | |
|---|---|---|---|---|
| | **RMSE** | **Coverage** | **RMSE** | **Coverage** |
| IBCF | 1.0814 | 20.57% | 0.8828 | 57.36% |
| UBCF | 0.5776 | 24.15% | 0.8140 | 59.81% |
| SVD | 0.8303 | 30.00% | 0.9093 | 64.15% |
| BMF | 1.1947 | 100.00% | 1.0214 | 100.00% |

The cluster-boosting leads to an alignment of the results, i.e. the RMSE increases for the UBCF and SVD, i.e. the two approaches with the lowest RMSE and in return, the RMSE decreases for the IBCF and BMF. However, using the cluster-boosted approach, the coverage is at least two (SVD) up to almost three (IBCF) times higher compared to the respective un-boosted approach with BMF as exception that holds a coverage of 100% due to its implementation.

## 5 Conclusion

We borrowed the technique of calculating higher-order co-occurrences from corpus linguistics and applied this technique to user sessions in order to group semantically similar learning resources without considering their content. We used the MACE data set to investigate the assumption that - similar to words - the higher order co-occurrence classes become stable and the learning resources within the classes are homogeneous. The intrinsic evaluation results clearly indicate that independent from the used association measure the resources within the clusters are similar according to their content and the clusters are well isolated. Though, the selection of the association measure has a large impact on the cluster size distributions. For example, using the association measure MI, the resulting clusters are mostly small and very dense, additionally, one large rag bag is created. On the contrary, the usage of cor-$\chi^2$ causes the loosest clusters, however, the rag bag is rather small, i.e. most resources are clustered. Thus, it depends on the application which measure should be preferred.

The application scenarios for the clusterings are manifold, e.g. to group resources returned by a search query or to enable a cluster-based browsing of a web portal's content. In this paper, the suitability of the clusterings to predict ratings and to enhance recommender systems is investigated. The evaluation of the rating predictions shows that the sole usage of the clusterings can even outperform the usage of well-known CF approaches. For example, the clustering-based recommendation utilising MI creates by far the most accurate ratings. Though it can only predict ratings for a subset of the user-item pairs in the test set the same holds true for most CF approaches. The other association measures create clusterings that can be used to predict half of the missing ratings with a much lower error rate than the only CF approach that predicts more ratings. When applying the cluster-boosted approach the strength of the different techniques can be combined. This results in an increased coverage and even the error rate can be decreased for some CF approaches.

Concluding, we emphasise that the clusterings created by calculating higher-order co-occurrences provide a real value. We plan to test this approach on more data sets and to integrate it in a running system to support teachers and learners.

## References


1. Verbert, K., Drachsler, H., Manouselis, N.: Dataset-driven research for improving tel recommender systems. In: 1st ACM International Conference on Learning Analytics and Knowledge, pp. 44–53 (2011)
2. Heyer, G., Quasthof, U., Wittg, T.: Text Mining: Wissensrohstoff Text. Konzepte, Algorithmen, Ergebnisse (2006)
3. Mahn, M., Biemann, C.: Tuning Co-occurrences of Higher Orders for Generating Ontology Extension Candidates. In: ICML Workshop on Learning and Extending Lexical Ontologies by Using Machine Learning Methods, pp. 28–32 (2005)
4. Hussain, S.F., Bisson, G.: Text Categorization Using Word Similarities Based on Higher Order Co-occurrences. In: Society for Industrial and Applied Mathematics International Conference on Data Mining (2010)
5. 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)
6. Schmitz, H.C., Wolpers, M., Kirschenmann, U., Niemann, K.: Contextualized Attention Metadata. In: Roda, C. (ed.) Human Attention in Digital Environments. Cambridge University Press, Cambridge (2011)
7. Evert, S.: Corpora and collocations. In: Lüdeling, A., Kytö, M. (eds.) Corpus Linguistics: An International Handbook. Mouton de Gruyter, Berlin (2008)
8. Amigó, E., Gonzalo, J., Artiles, J., Verdejo, F.: A comparison of extrinsic clustering evaluation metrics based on formal constraints. Information Retrieval 12(4), 461–486 (2008)
9. Davies, D.L., Bouldin, D.W.: A Cluster Separation Measure. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 1(2), 224–227 (1979)
10. Raskutti, B., Leckie, C., Rd, B.: An Evaluation of Criteria for Measuring the Quality of Clusters. In: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, pp. 905–910 (1999)
11. Vries, C.M.D., Geva, S., Trotman, A.: Document Clustering Evaluation: Divergence from a Random Baseline. The Computing Research Repository (2010)
12. 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)
13. Lee, J., Sun, M., Lebanon, G.: A Comparative Study of Collaborative Filtering Algorithms. ArXiv e-prints (May 2012)

# Do Optional Activities Matter in Virtual Learning Environments?

José A. Ruipérez-Valiente[1,2], Pedro J. Muñoz-Merino[1], Carlos Delgado Kloos[1], Katja Niemann[3], and Maren Scheffel[4]

[1] Universidad Carlos III de Madrid, Avenida Universidad 30, 28911 Leganés, Madrid, Spain
{jruipere,pedmume,cdk}@it.uc3m.es
[2] IMDEA Networks Institute, Av. del Mar Mediterráneo 22, 28918 Leganés, Madrid, Spain
[3] Fraunhofer Institute for Applied Information Technology (FIT), Schloss Birlinghoven, 53754 Sankt Augustin, Germany
katja.niemann@fit.fraunhofer.de
[4] Open Universiteit Nederland, Valkenburgerweg 177, 6419 AT Heerlen, The Netherlands
maren.scheffel@ou.nl

**Abstract.** Virtual Learning Environments (VLEs) provide studentts with activities to improve their learning (e.g., reading texts, watching videos or solving exercises). But VLEs usually also provide optional activities (e.g., changing an avatar profile or setting goals). Some of these have a connection with the learning process, but are not directly devoted to learning concepts (e.g., setting goals). Few works have dealt with the use of optional activities and the relationships between these activities and other metrics in VLEs. This paper analyzes the use of optional activities at different levels in a specific case study with 291 students from three courses (physics, chemistry and mathematics) using the Khan Academy platform. The level of use of the different types of optional activities is analyzed and compared to that of learning activities. In addition, the relationship between the usage of optional activities and different student behaviors and learning metrics is presented.

**Keywords:** optional activities, Khan Academy, learning analytics, MOOCs.

## 1 Introduction

Virtual Learning Environments (VLEs) usually provide learning activities (e.g., watching videos or solving exercises) to improve student learning. But these platforms usually also contain functionalities and tools that are optional, i.e., they are not mandatory to successfully finish a learning course. In this work, we consider optional activities as functionalities that are present in different learning platforms, but that are not directly related to performing an activity to acquire learning. In some cases, such as our case study, teachers do not inform students about the optional activities available. Instead, students usually discover these functionalities during their interactions with the system. Some examples of optional activities are the setting of the students' own profiles or the setting of their own goals.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 331–344, 2014.
© Springer International Publishing Switzerland 2014

Although there are usually many types of optional activities in VLEs, research studies usually focus on the analysis of learning activities but not on the optional activities. The study reported by Dyckhoff et al. [1] shows a list of learning analytics indicators used in the literature and its categorization. We can see that most of the indicators presented by Dyckhoff et al. [1] are related to learning activities instead of optional activities. The analysis of optional activities enables teachers to know which optional activities are used more, which of them need to be promoted, whether there is a relationship with the use and correct solving of learning activities, or their relationship with other student behaviors.

In this research, we analyze the use of optional activities in a real experiment using the Khan Academy[1] platform with 291 students interacting in three different courses. The objectives of this analysis are the following:

- Knowing the level of use of the different types of optional activities, giving the percentage distribution between them, but also with respect to the learning activities.
- Relating the use of optional activities to the correct solving of learning activities.
- Relating the use of optional activities to other student behavior metrics and parameters such as hint abuse, hint avoidance, total time, progress in the platform, etc.
- Checking for categorical associations between the use of the different optional activities and also with other variables such as gender.

## 2 Related Work

The current common use of VLEs in different learning contexts presents many opportunities to analyze the interaction of students in these environments. These analyses are usually focused on learning activities but not on optional ones. A previous work [2] has analyzed which of the available tools and functionalities are more important than others in Moodle[2] and .LRN[3]. The results of that work indicate that the most-rated functionalities are submission management systems with the teacher's feedback, visualizations of qualifications, assessments, file downloads, FAQs and forums. However, it has also been shown that students do not always make adequate choices in the learning tools they can use [3]. In addition, the use of tools seems to be influenced by the student's characteristics, the kind of tool being considered, and the advice students receive.

One of the most widely used learning activities in e-learning environments is the use of forums to support social interaction; this provides a good opportunity to analyze social interactions in e-learning platforms. For instance, Rabbany et al. [4] analyzed the importance of social network analysis for mining structural data and how this is applicable to the student relationships in e-learning environments. Other works have statistically analyzed forum activity to draw conclusions such as the effect of the

---

[1] https://www.khanacademy.org/
[2] https://moodle.org/
[3] http://dotlrn.org/

teaching staff's participation or the evolution of forum discussions throughout the duration of the course [5]. Other analyses dealt with the adding of an electronic voting system (EVS) to regular courses [6, 7]. Each type of platform has different types of features to be analyzed. For example the analysis of the use of different types of hints in an Intelligent Tutoring System (ITS) [8]. There are other works that have analyzed the use and outcome of different implementations of gamification in e-learning, e.g., a reward system in the Blackboard platform [9] or a reputation system in an edX Massive Open Online Course (MOOC) [10]. The interaction of students in a Moodle course with wiki articles (where students can peer review each article), discussion forums and also access to optional material such as scientific articles or lecture videos have also been looked at [11]. In addition, many studies inspected the relationship between forum participation and learning outcomes: forum activity, for example, in UNED COMA courses is a useful indicator for success [12]. However, it was also discovered that online participation and interaction does not necessarily translate into higher grades [13].

The use of the different tools commented on in the previous paragraph dealt with learning activities. On the other hand, there are also other works that addressed the use of optional activities. Some studies have dealt with the relationship between student interactions with a specific e-learning motivation [14]. It was found that interactions such as uploading a personal profile photo (an optional activity) and participating in forums are positively correlated to motivation. Moreover, findings suggest that the benefits from optional forum assignments are very low whereas compulsory forum assignments might be related to higher performance [15]. In addition, the work in [16] made an in-depth analysis of the use of video annotation, which can be seen as an optional activity for students.

As a conclusion, there has been extensive work to analyze the use of learning activities and relate it to different indicators such as learning outcomes. However, there has not been many works focused on optional activities. As a difference of this work with respect to the presented works about optional activities, this paper gives details of the total use distribution of several different optional activities, relates this use to exercise correct solving indicators, as well as to some student behaviors such as some used by researchers in previous works [17].

## 3 Measuring Optional Activities in VLEs

In this section we propose a simple general measure to take into account the different optional activities that each student has used.

$$optional\ activities = \frac{100}{max_{points}} \sum_{i=1}^{N} w_i * points_i$$

The general formula hereby proposed for optional activities is specified above, where $N$ is the number of different optional activities, $points_i$ is the number of points acquired by the student in the optional activity $i$, $w_i$ is the weight of the activity $i$ and $max_{points}$ is the maximum number of points that can be achieved in the whole course if you have made the maximum use of all the optional activities (taking also into account weights). The points are assigned by researchers but it does not mean that the platform assigns points automatically. We should take into account the following premises in order to adapt it to each context:

- This is a general formula that can be configured to each context with a measure ranging from 0 (i.e., a student who has not used any optional activity) to 100 (i.e., a student who has made the maximum use of all the optional activities).
- The amount of points that can be achieved by completing each optional activity is a configuration parameter that should be analyzed for each context. A student can either earn several points if he has used the same activity more than once, or just earn one point no matter the number of times the student has used it. The model gives a general way of measuring it in different contexts.
- Finally, each activity can have a different weight, taking into account the importance within the set of optional activities that are being considered in the case study. For example, it could be regarded as more important to complete learning goals or participate in forum activities than to update your profile avatar, thus the weight should be different for that case.

### 3.1 The Application to Khan Academy

In this sub-section we analyze a particular platform, i.e., Khan Academy, as it was used for pre-graduate courses at Universidad Carlos III de Madrid. Section 4.1 describes this experiment in more depth, while this section describes the optional activities we have taken into account in the Khan Academy courses. We have divided these activities in two groups: optional activities that are related to learning and others that are not related to learning. First, the activities that are related to learning include:

- Feedback: Comments that students post to videos of the course are considered as feedback.
- Votes: Students can vote down (-1), be indifferent to (0) or vote up (+1) the feedback that other students have posted to videos. Figure 1 shows an example of a comment that has some votes. The name and message of the author have been blurred to preserve anonymity.
- Goal: Students can set goals, i.e., they choose a selection of videos or exercises that must be completed, and when they finish the goal they obtain an additional amount of points. Figure 2 shows an example about how to set a custom goal.

On the other hand, we have taken into account other optional activities that are not related to learning. These activities come from social networks and games environments:

- Profile avatar: Students can change the default avatar of their profile. They have basic access to different avatar images, and they can earn more by acquiring points in their interaction with the platform functionalities.

- Badge display: Students can personalize a selection of badges to be displayed in their personal profile. The badges that can be displayed are the ones that each student has earned previously. Figure 3 shows a portion of the personal profile where the profile avatar and the badge display can be observed.

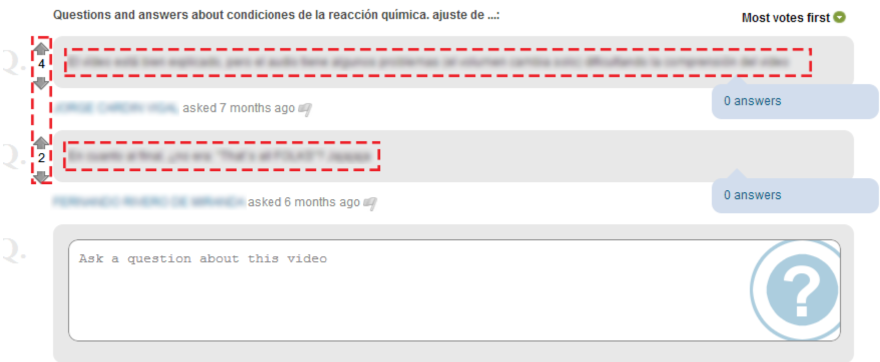



**Fig. 1.** Students' feedback and votes in a video example can be observed inside the red dotted line. The authors and comments have been blurred to preserve anonymity.

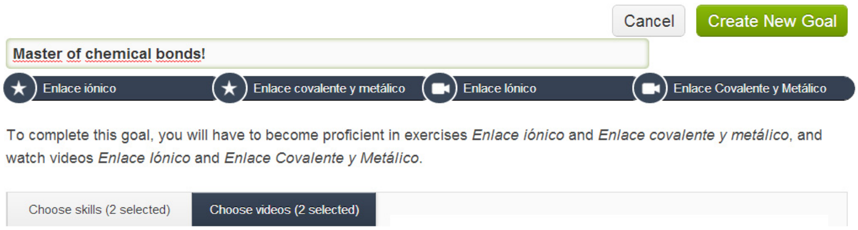



**Fig. 2.** Custom goal example. The student has to master chemical bonds exercises and videos.

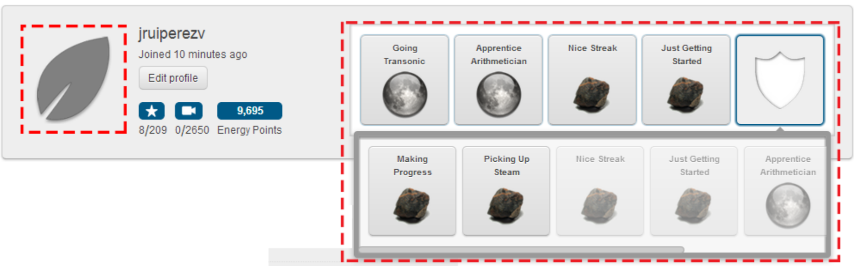



**Fig. 3.** Personal profile of a student in our Khan Academy instance. The default avatar image and badge display are pointed out by the red dotted line.

For our case study we wanted to measure which optional activities have been used, so our score methodology is adapted to this purpose. We decided that all the activities have the same weight (although goals can provide a maximum of two points, as explained later). We do not wish to give more importance to any of the activities in this case study, just to check which of the activities have been used. Specifically, the scoring method is designed as follows: A user gets one point whenever he selects an avatar image, selects at least one badge to be displayed, writes a feedback about a video, votes for any of the comments of other students or starts at least one goal (an additional point is given when the user finishes at least one started goal).

Consequently, there are five different activities, and students can earn a maximum quantity of six points if they do them all. The general metric to measure the use of optional activities is adapted to this specific case studio in the next line:

$$optional\ activities = \frac{100}{6}\sum_{i=1}^{5} points_i$$

## 4 Case Study

In this section we present the case study about the usage of optional activities. The first sub-section describes the context where the experiment was conducted. The second sub-section offers an overview of the total use of the optional activities within the setting, as well as a comparison between the optional activities usage and regular courseware. The last sub-section analyzes the relationship between the use of optional activities and other metrics that we have defined in others works.

### 4.1 Description of the Experiment

Universidad Carlos III de Madrid uses the Khan Academy platform to apply an "inverted classroom" methodology [18], in which students learn and review concepts by using the provided Khan Academy courses at home and then take the face to face lessons in the university. These courses are for freshmen students who are entering a science degree in order to review the concepts required during their first year of university. We used the data generated in three courses (physics, chemistry and mathematics) that took place in August 2013. These courses are composed of sets of videos and exercises, which have been designed by the instructors. Although it was not mandatory for the students to access these courses, it was strongly recommended; this is an important fact when measuring the use of the platform. It is also noteworthy that students were not informed of the optional activities available.

Most of the participants were freshmen around 18 years old. The number of students was different for each course, and some students had to take more than one course (depending on the science degree they were entering). There were 167 students in the physics course, 73 in chemistry and 243 in mathematics; this is a total number of 483 data samples. However, the total number of unique students was 291, as some students were enrolled in more than one course.

### 4.2 Use of Optional Activities

This section includes a quantitative analysis of the use of the optional activities. This usage is also compared to the participation ratios with the regular items (such as exercises and videos) of the courses. We provide data from each of the courses separately and also for the overall results of all courses. Figure 4 shows results of the optional activities use for each activity and class. Each bar represents the percentage of students in the class who have used the activity on the left axis. In addition the last metric provides the percentage of users who have used at least one of the activities. Each one of the courses is represented by a different color where blue stands for chemistry, dark grey for physics, green for mathematics and red for the students in all classes.

The results shown in figure 4 take into account all the students who logged in at least once to the Khan Academy platform. Consequently, it is certain that many of these students did not interact much with the system, neither with optional activities nor with learning activities. From figure 4 we can extract the following main conclusions. The optional activities used the most are the configuration of a profile avatar and the badge display. Although the exact percentage numbers differ from one course to another, on average we can draw the following results; with 10.8% (avatar) and 12% (badges) respectively for all courses, they are by far the most used optional activities in all courses. A possible reason for this could be that these students, who are aged around 17–19 years, are comfortable using activities originating from a social network or gaming context.

On the other hand, optional activities that are related to learning (feedback, vote and goal) have been used much less (4.1%, 6.6% and 6.2%, respectively) in all courses. The activity which has been used the least is feedback. A reasonable argument is that writing a feedback answer about a video generally requires a greater effort than just simply changing an avatar, for example. Furthermore, Moodle forums were also enabled for students during these courses, and they conducted most of their social interaction in these forums.

Finally, the 23.2% of the students of all courses who logged in at least once on the platform used at least one of the five optional activities considered in this study. In addition, results show a difference in the use of optional activities between the three courses. The chemistry course has the highest ratio of students who used at least one optional activity (30.1%), whereas physics has the lowest ratio (18%). Further research would be needed to establish possible reasons for these differences. It is important to remember that students did not have knowledge of the optional activities available; they were only informed of the courseware. This can be one of the main reasons for the low use ratios for these activities. However, taking into account that these courses were not mandatory, and that these activities were not announced, the usage ratios are still low. It is also interesting to check on more specific details such as the ratio of finished goals and the types of votes. A total number of 55 goals were started taking into consideration all courses, and 28 of them (50.9%) were finished. These are high finishing ratios taking into account that goals are optional activities and also related to the learning process. We have computed 40 votes in all courses, 26 of them were positive (65%), 13 of them were indifferent (32.5%) and only one of them was negative (2.5%). These results indicate that most of the users vote for positive reasons given these conditions and it is very unlikely that they vote negatively on others students.

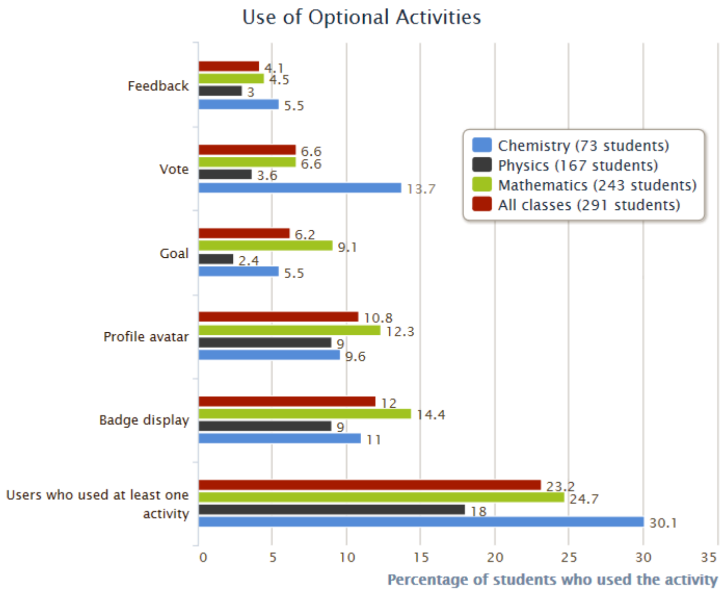



**Fig. 4.** This chart represents the percentage of users who have used each of the activities in each course

We can establish a comparison with the access to the regular learning activities such as exercises and videos and also the use of optional activities. This comparison can be seen in Table 1. The purpose of this table is to give a sense of how much students have used the regular activities in comparison to optional activities. We have divided the use of regular and optional activities in five intervals and we show the percentage of students from all courses in each interval. It is noteworthy that we have a total number of 483 student samples because some of the students participated in several courses and these statistics take into account all the samples. The first thing to notice is that only 12 students (2.48%) who logged in on the platform did not use any of the regular learning activities whereas 371 of the students (76.81%) did not use any of the optional activities. This is a huge difference that already gives insight about the low use of optional activities compared to the use of regular learning activities. On the other end we can notice that 19 students (3.93%) used all the regular learning activities while only one of the students (0.21%) used all the optional activities. These results are even more striking keeping in mind that the amount of learning activities is above 40 in all courses whereas the number of optional activities taken into account in the study are only five. Finally, we can check that the use of activities in the 1-99% interval declines gradually and it is always superior for the regular activities.

**Table 1.** Comparison between the use of regular learning activities versus the use of optional activities. Each cell represents the percentage of students for each interval.

| Type of activity | Percentage of activities accessed | | | | |
|---|---|---|---|---|---|
| | 0% | 1-33 % | 34-66% | 67-99% | 100% |
| **Regular learning activities** | 2.48% | 51.55% | 23.19% | 18.84% | 3.93% |
| **Optional activities** | 76.81% | 18.43% | 4.14% | 0.41% | 0.21% |

### 4.3 Relationships between Different Metrics

In this sub-section we present different relationships between the usage of the optional activities (using the metric formula specified in section 3.1.) with other metrics that we have calculated in other works [19]. The metrics are the following: exercise and video access, exercise and video abandonment, total time spent in exercises and videos, following of recommendations, hint avoidance (not solving an exercise correctly but not asking for hints), hint abuse (asking for too many hints without reflecting on previous ones), video avoidance (not solving an exercise correctly but not watching a related video) and unreflective user.

Table 2 shows the Pearson Correlation (N = 291, two-tailed significance) of the optional activities measured with each of the aforementioned metrics. The correlations that are significant at the 99% level are marked with an asterisk. The data shows that the most significant correlations are with the total time (0.491), and also with the percentage of accessed exercises and videos (0.429 and 0.419). In addition, another significant but negative correlation is with exercise and video abandonment (-0.259 and -0.155). This negative correlation means that users who abandon exercises and videos use less optional activities than others, which makes sense because probably they have spent less time in the environment than others. Finally, the results indicate that the behavioral metrics (following recommendations, video and hint avoidance, hint abuse and unreflective user) are not significantly correlated to the use of optional activities. We found that there is no relationship between the follow recommendations profile and using the optional activities, although we initially thought that there could exist a relationship due to the fact that the use of optional items can be regarded as an exploring behavior. It is also intriguing to check how the use of optional activities is related to correct solving metrics; we use the percentage of proficient exercises, which are the exercises that students have mastered at the maximum level in Khan Academy. Table 3 shows the Pearson Correlation between the percentage of proficient exercises with the global measure of optional activities and each optional activity separately. The results show that the use of optional activities is significantly correlated with the percentage of proficient exercises, which could mean that it is also related to learning. The most significant correlation (0.553) is with the global measure of optional activities; this relationship points out that the use of optional activities might be used as an indicator to know how much students have progressed in the exercises. Avatar and display badge (0.415 and 0.418) are the optional activities that have been most highly correlated with the percentage of proficient exercises, whereas feedback and vote (0.205 and 0.243) have been the least. This might be surprising because feedback and vote are supposed to be related to the learning process, thus that might invite to think that they should be more highly correlated with resolving exercises correctly than avatar and display badges which are not related to the learning process.

**Table 2.** Bivariate Pearson correlation of optional activities with others metrics. (*): Correlation is significant at the 0.01 level (2-tailed).

| | Exercises accessed: | Videos accessed: | Exercise abandonment: | Video abandonment: | Total time: |
|---|---|---|---|---|---|
| **Optional activities sig. (2-tailed) N = 291** | 0.429* (p=0.000) | 0.419* (p=0.000) | -0.259* (p=0.000) | -0.155* (p=0.008) | 0.491* (p=0.000) |
| | **Hint abuse:** | **Hint avoider:** | **Follow recommendations:** | **Unreflective user:** | **Video avoider:** |
| | 0.089 (p=0.131) | 0.053 (p=0.370) | -0.002 (p=0.972) | 0.039 (p=0.507) | -0.051 (p=0.384) |

**Table 3.** Bivariate Pearson correlation of the percentage of proficient exercises with the different optional activities

| | Optional activities: | Goal: | Feedback: |
|---|---|---|---|
| **Proficient exercises sig. (2-tailed) N = 291** | 0.553* (p=0.000) | 0.384* (p=0.000) | 0.205* (p=0.000) |
| | **Vote:** | **Avatar:** | **Display badges:** |
| | 0.243* (p=0.000) | 0.415* (p=0.000) | 0.418* (p=0.000) |

We can also make a comparison with other categorical variables by cross-tabulating the different results. The categorical variables we have used in this analysis are gender, the course and the separate use of each optional activity defined as ‘yes’ or ‘no’ per each student. We can analyze the relationship between these variables by using cross-tabulation techniques that are also known as contingency tables. Whether the established relation is really significant can be calculated by applying the Pearson Chi-Square Test for categorical data. If the expected count assumptions (the expected count of each cell must be above five) of the Pearson Chi-Square are not met, we can apply the Fisher’s Exact Test.

The first cross tabulation is established between gender and the use of the different optional activities. Results reveal that women more often use goals, avatar and badge display whereas men use feedback and vote activities more often. However, the

Pearson Chi-Square Test shows that the only significant relationship is the one with feedback use. The test indicates (with a value of 2.80, p = 0.048) that it is statistically significant that men use the feedback activity more than women; the minimum expected count in each cell is higher than five, so the Pearson Chi-Square Test assumption is fulfilled for this case. The participation differences in web-based learning environments by gender have also been addressed in other works [20].

The second analysis takes the course and the use of the different optional activities into account. The results reveal that chemistry is the course where feedback and votes were used the most whereas mathematics is the course in which goals, avatar and badge display were used the most. Furthermore, as already mentioned in the last subsection, physics is the course that made the least use of optional activities. In order to know which of the results are significant, we applied the Fisher's Exact Test whose assumptions are met (as the Pearson Chi-Square Test assumption of a minimum count of five is not given). The test shows (6.58, p = 0.034) that the relationship between course and goal is significant, so it is statistically significant that mathematics is the course where goal-activity is used most. In addition, the relationship between the use of votes and the course is also significant (7.74, p = 0.019), where chemistry is the course that used votes the most.

The last analysis carried out is that between all the categorical variables that represent the use of each optional activity. To this end we chose a log linear analysis which allows the comparison of three or more categorical variables in order to determine if there is an association between two or more of them. The factors of the test are the use of each optional activity separately (yes or no) for each student. Table 4 shows the cell count of a log linear analysis of only those associations where the observed count is above or equal to 1% of the cases.

**Table 4.** Cell count of the most important associations of the log linear analysis between all optional activity categorical variables

| **Used goal?** | **Used feed-back?** | **Used vote?** | **Used avatar?** | **Used display badges?** | **Observed** | |
|---|---|---|---|---|---|---|
| | | | | | Count | % |
| No | No | No | No | No | 371 | 76.8% |
| No | No | No | No | Yes | 20 | 4.1% |
| No | No | No | Yes | No | 14 | 2.9% |
| No | No | No | Yes | Yes | 14 | 2.9% |
| No | No | Yes | No | No | 15 | 3.1% |
| No | Yes | No | No | No | 8 | 1.7% |
| Yes | No | No | No | No | 5 | 1.0% |
| Yes | No | No | Yes | Yes | 7 | 1.4% |

Table 4 allows us to see which ones are the most typical associations in percentage. The higher counts are the use of display badge (4.1%), the use of avatar (2.9%), the use of both display badge and avatar (2.9%) and the use of votes (3.1%). The data indicate that there are probably underlying associations between the use of these activities, consequently we check other tests to see if it is really significant. The z-score values show that the most significant relationships are between the use of avatar and display badges ($z = 2.68$, $p = 0.007$), between the use of feedback and votes ($z = 2.26$, $p = 0.008$) and also between the use of goal and avatar ($z = 2.1$, $p = 0.036$). These results make sense because an association between the use of avatar and display badge is related to activities that come from customizing your personal profile, and the association between the use of feedback and votes are activities related to participation in a forum. In addition, there is a three-way significant relationship between the use of goals, avatar and display badge ($z = 1.96$, $p = 0.05$), which is also interesting because these three activities are related to gaming or social networks environments.

## 5 Conclusions and Future Work

In this work, we analyze the use of optional activities in a VLE using data from real experiments in the Khan Academy platform with 291 students from three different courses. Results indicate that the use of the optional activities under the conditions of this experiment (the use of the platform was not mandatory and instructors did not inform their students about the optional activities) has been very low. Therefore, we would recommend telling the students about the availability of optional activities or tools for future experiments.

Additionally, results show that the optional activities that were used the most are not related to learning (avatar and display badges), so we would recommend instructors to encourage students to use optional activities related to learning such as feedback, votes, or goals. Another interesting finding is that more than half of the goals that were started, were finished by the students; we think this type of optional activity is good to improve the engagement of students to the learning process.

We have provided a detailed analysis of the relationship between the use of optional activities and several student behaviors as well as student metrics of use of the platform. This analysis indicates that the use of optional activities is significantly related to the total time spent in the platform and the progress in exercises and videos. Another important result is that none of the behavioral metrics is correlated with the use of optional activities.

In order to answer the title question, the use of optional activities did not matter a lot for students under the conditions of this experiment in the sense that these optional activities were under-used. However, we think that the use of optional activities does matter and their use should be promoted. The results have shown that the use of optional activities is significantly correlated to the total time and use of regular learning activities. Also, optional activities have been found to correlate with solving exercises metrics. Although we cannot infer that the use of optional activities is the cause of better learning, this hypothesis remains open for future research. In addition the categorical variable analysis revealed that there are several statistically significant co-occurrence associations; for example between the use of both display badge and avatar which come from games environments, and also the use of both feedback and votes which are related to forum activities. This could be used to cluster students or divide them in groups with similar preferences.

However, despite several significant results, this is a preliminary research and we still are not prepared to confirm the causes of some results, and more work is required to answer those questions. An important step for future work would be to understand the relationship between the use of optional activities and learning achievement. This way we could know what optional activities should be reinforced for future experiments. Finally, in the future we would like to review the optional activities available in other VLEs and formulate a common framework.

**Acknowledgements.** Work partially funded by the EEE project, "Plan Nacional de I+D+I TIN2011-28308-C03-01" and the "Emadrid: Investigación y desarrollo de tecnologías para el e-learning en la Comunidad de Madrid" project (S2009/TIC-1650)". Maren Scheffel's contribution is funded by the European Union under FP7 Grant Agreement No 317620 (LinkedUp) and No 619424 (LACE).

## References

1. Dyckhoff, A.L., Lukarov, V., Muslim, A., Chatti, M.A., Schroeder, U.: Supporting Action Research with Learning Analytics. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge, pp. 220–229. ACM, New York (2013)
2. Muñoz-Merino, P.J., Kloos, C.D., Seepold, R., García, R.M.C.: Rating the importance of different LMS functionalities. In: 36th Annual Conference on Frontiers in Education Conference, pp. 13–18. IEEE, San Diego (2006)
3. Clarebout, G., Elen, J.: Tool use in computer-based learning environments: Towards a research framework. Computers in Human Behavior 22(3), 389–411 (2006)
4. Rabbany, R., Elatia, S., Takaffoli, M., Zaïane, O.R.: Collaborative learning of students in online discussion forums: A social network analysis perspective. In: Educational Data Mining. SCI, vol. 524, pp. 1–25. Heidelberg, Germany (2014)
5. Brinton, C.G., Lam, H., Ming, F., Wong, F.M.F.: Learning about social learning in MOOCs: From statistical analysis to generative model. CoRR, abs/1312.2159 (2013)
6. Cutrim, E.S.: Using a voting system in conjunction with interactive whiteboard technology to enhance learning in the English language classroom. Computers & Education 50(1), 338–356 (2008)
7. Kennedy, G.E., Cutts, Q.I.: The association between students' use of an electronic voting system and their learning outcomes. Journal of Computer Assisted Learning 21(4), 260–268 (2005)
8. Zhou, Y., Freedman, R., Glass, M., Michael, J.A., Rovick, A.A., Evens, M.W.: Delivering hints in a dialogue-based intelligent tutoring system. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI 199), pp. 1–7. American Association for Artificial Intelligence, Menlo Park (1998)
9. 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)
10. Coetzee, D., Fox, A., Hearst, M.A., Hartmann, B.: Should your MOOC forum use a reputation system? In: Proceedings of the 17th ACM Conference on Computer Supported Co-

operative Work & Social Computing (CSCW 2014), pp. 1176–1187. ACM, New York (2014)
11. Hecking, T., Ziebarth, S., Hoppe, H.U.: Analysis of dynamic resource access patterns in a blended learning course. In: Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, pp. 173–182. ACM, New York (2014)
12. Santos, J.L., Klerkx, J., Duval, E., Gago, D., Rodríguez, L.: Success, activity and drop-outs in MOOCs an exploratory study on the UNED COMA courses. In: Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, pp. 98–102. ACM, New York (2014)
13. Davies, J., Graff, M.: Performance in learning: Online participation and student grades. British Journal of Educational Technology 36(4), 657–663 (2005)
14. Munoz-Organero, M., Munoz-Merino, P.J., Delgado Kloos, C.: Student behavior and interaction patterns with an LMS as motivation predictors in e-learning settings. IEEE Transactions on Education 53(3), 463–470 (2010)
15. Gafni, R., Geri, N.: The value of collaborative e-learning: Compulsory versus optional online forum assignments. Interdisciplinary Journal of E-Learning and Learning Objects 6, 335–343 (2010)
16. Gasevic, D., Mirriahi, N., Dawson, S.: Analytics of the Effects of Video Use and Instruction to Support Reflective Learning. In: Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, pp. 123–132. ACM, New York (2014)
17. Aleven, V., McLaren, B.M., Roll, I., Koedinger, K.R.: Toward Meta-cognitive Tutoring: A Model of Help Seeking with a Cognitive Tutor. International Journal of Artificial Intelligence in Education 16(2), 101–128 (2006)
18. Lage, M.J., Platt, G.J., Treglia, M.: Inverting the classroom: A gateway to creating an inclusive learning environment. The Journal of Economic Education 3(1), 30–43 (2000)
19. Muñoz-Merino, P.J., Ruipérez, J.A., Delgado, C.: Inferring higher level learning information from low level data for the Khan Academy platform. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge, pp. 112–116. ACM, New York (2013)
20. Caspi, A., Chajut, E., Saporta, K.: Participation in class and in online discussions: Gender differences. Computers & Education 50(3), 718–724 (2008)

# Reconstructing IMS LD Units of Learning from Event Logs

Juan C. Vidal, Manuel Lama, Borja Vázquez, and Manuel Mucientes

Research Center on Information Technology (CiTIUS),
University of Santiago de Compostela
{juan.vidal,manuel.lama,borja.vazquez,manuel.mucientes}@usc.es

**Abstract.** In this paper a novel approach to facilitate the reuse of units of learning (UoLs) is presented. Typically, e-learning platforms do not provide the means to retrieve designed UoLs in a standardized format to be reused in a different platform, but they have in common that the students and teachers interaction with the system is logged to files. Taking this into account, we propose to use these logs and apply a three steps re-engineering approach to translate these UoLs into an accepted educational modelling language, specifically IMS LD. In the first step, the sequence of activities and their functional dependencies are learned by a process mining algorithm. In the second step, another algorithm analyses the variables and their value change in order to learn the adaptation rules that may have been defined in the UoL. And finally, in the last step the inferred process structure and rules are matched with the typical structure of activities, acts, and plays defined by IMS LD.

**Keywords:** IMS LD, process mining, adaptive rules mining, learning analytics, learning flows reengineering.

## 1 Introduction

In the last decade, an important effort to develop Educational Modelling Languages (EMLs) has been made. The aim of these languages is to describe from a pedagogic point of view the learning design of a course: that is, the flow of learning activities undertaken by the learners to achieve the objectives of a course using educational content and services. From these EMLs, the IMS Learning Design specification [3] has emerged as the *de facto* standard for representing learning designs that can be based on a wide range of pedagogical techniques. Although there is some controversy about whether IMS LD is too complex to be understood by teachers from a practical point of view [5], most of authors highlights this complexity as a barrier for the adoption of IMS LD [8].

To deal with this issue a number of user-friendly authoring tools have appeared. These tools provide graphical interfaces that allow to hide the complexity of the IMS LD control structure and adaptive components [7]; offer visual templates or patterns for creating learning designs based on pedagogical strategies such as collaborative learning or project based learning [6]; or translate the

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 345–358, 2014.
© Springer International Publishing Switzerland 2014

learning design to the IMS LD specification from authoring tools that are not based on IMS LD [4]. However, even with these tools the authoring process of IMS LD units of learning (UoLs) is not easy for teachers when the UoLs are complex or require to use advanced features of IMS LD.

The automatic reconstruction of UoLs could relieve this issue. In [8] authors present an approach that focus on the reegineering of IMS LD UoLs based on a visual language, which hides the complexity of the model. On the other hand, this approach also requires a close collaboration between developers and teachers to simplify the gap between the technical and pedagogical point of view of the UoL. The main drawback of this approach is that the UoL is not automatically reconstructed from the scratch and needs the supervision of teachers and even developers. Other similar approaches have been proposed in the literature [2,11], although not for IMS LD.

In this paper, we present an approach to *automatically* reconstruct the IMS LD representation of a UoL from the events generated by the learners in the virtual learning environment. This objective is achieved in three different steps. Firstly, the learning flow of the UoL is automatically extracted through a *process discovery algorithm* [1] which guarantees the completeness of the discovered learning flow. Then an algorithm based on the knowledge about the IMS LD control structure is applied to determine which IMS LD components should be created. Finally, the adaptive rules of the UoL are automatically extracted from the event logs by a *decision tree* learning algorithm and integrated in the IMS LD structure.

The paper is structured as follows: in Section 2 we present the framework that supports the mining of log files and the reconstruction of IMS LD. In sections 3 and 4 we precisely detail the how the learning flow and adaptive rules are mined from the logs, respectively, while in Section 5 we close the circle and describe how we transform these two models to IMS LD. Finally, in Section 6 we present the conclusions of the work.

## 2 Framework

Fig. 1 depicts the conceptual framework for reconstructing IMS LD UoL from the events generated by learners. The framework is composed by the following components:

- *Educational World.* Teachers and learners are the typical participants in any Learning activity. On the one hand, *teachers* design the learning flows based on some educational methodology, and support the learning activities of the course. On the other hand, *learners* are the core of the educational world since they undertake the learning flow activities by using the resources and services available in the virtual learning environment.
- *Virtual Learning Environment.* From an educational point of view, virtual learning environments (VLEs) provide the means to carry out the learning activities planned for a UoL, allowing learners to access to the learning contents and executing the services required to facilitate those activities such

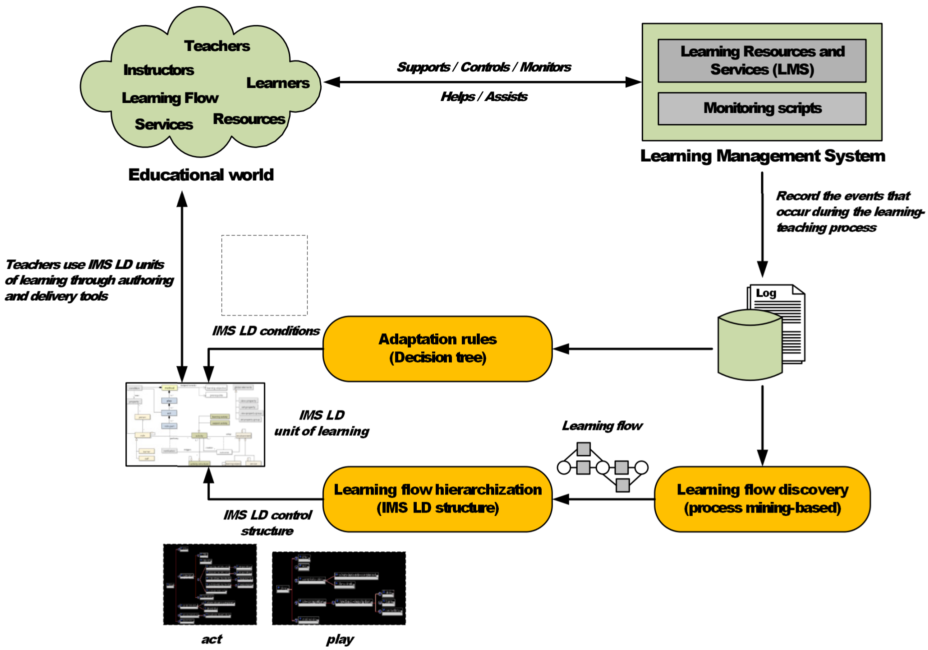



**Fig. 1.** Framework for IMS LD UoLs reconstruction

as interacting with other learners or looking for information in libraries. Furthermore, VLEs detect and register all the relevant events generated by learners when undertake the learning activities. These events are stored in an event log database that contains data about the activities execution, including who participates in an activity, when it has been performed, the properties values of the UoL such as a mark of a test, and so forth.

– *Learning Flow Discovery* This component implements the algorithm whose aim is to discover the workflow of learning activities, i.e., the learning flow, that learners undertake during a UoL. Note that this algorithm must guarantee that *all the learning paths* followed by learners are represented in the discovered learning flow. Furthermore, this discovered learning flow should be as simple as possible in order to facilitate the hierarchization of the learning flow into the IMS LD control structure.
– *Learning Flow Hierarchization.* Once the learning flow has been extracted from the event logs, an algorithm will translate this learning flow to the control structure defined by the IMS LD specification. This algorithm starts by detecting sequences and selections of learning activities and then do apply the knowledge about the IMS LD control structure to create activity structures, acts or plays.
– *Adaptation Rules.* The IMS LD specification has an extensive set of adaptation conditions, but this framework is focused on extracting the conditions related to the selection of learning activities (show and hide mechanism)

based on the changes in the properties values of the UoL. A decision tree technique was implemented to automatically obtain these conditions, since it is an effective approach to deal with this kind of problems.

In the following sections we will detail the last three components that constitute the core of this framework.

## 3 Mining the Learning Flow from Event Logs

The aim of the process discovery algorithm (PDA) [1] is to identify the workflow that represents the learning flow followed by the learners during a UoL. To achieve this objective, the PDA will only needs to process the metadata information provided by the event logs. From the perspective of process mining, the quality of a process discovery algorithm is measured taking into account the following metrics:

– *Completeness*, which indicates how much of the behaviour observed in the event log can be reproduced by the discovered learning flow. Thus, a discovered learning flow is complete when it can reproduce all the events contained in the log database. In order to guarantee correct reconstructions of IMS LD UoLs, all the activities undertaken by learners have to be included in the mined learning flow. Therefore, the completeness of discovered learning flows is a *hard requirement* for the PDA.
– *Precision*, which measures if the discovered learning flow is overly general, allowing an additional behaviour that is not represented in the log. Thus a discovered learning flow is precise when it cannot reproduce event traces that are not available in the log database. From the point of view of the IMS LD reconstruction to discover precise learning flows is not a requirement as hard as completeness, since the extra behaviour has not been undertaken by any learner.
– *Minimality*, which refers to discovered learning flows with the minimal structure that reflects the behaviour contained in the log database. A desirable requirement for the PDA is to obtain simple learning flows, since the simpler is the discovered learning flow, the easier is to reconstruct the IMS LD control structure.

Taking these measures into account, we have developed a genetic algorithm that discovers complete learning paths with very high values for precision and simplicity. Algorithm 1 describes this algorithm, where the first three steps correspond to its initialization with $t$ representing the number of iterations of the algorithm, *timesRun* is used to detect situations in which the search gets stuck, and restart counts the number of times the algorithm is reinitialized. The execution cycle of the genetic algorithm is as follows:

– A population is created with a group of individuals where each individual is a potential solution, i.e., a learning flow. In this algorithm, learning flows are

**ALGORITHM 1.** Genetic algorithm for discovery of learning flows

```
Initialize population
Evaluate population
t = 1, timesRun = initialTimesRun, restarts = 0
while t ≤ maxGenerations && restarts < maxRestarts do
    Selection
    Crossover
    Mutation
    Evaluate new individuals
    Replace population
    t = t + 1
    if bestInd(t) == bestInd(t − 1) then
        timesRun = timesRun − 1
    end
    if none of the individuals of the population have been replaced then
        timesRun = timesRun − 1
    end
    if timesRun < 0 then
        Reinitialize population
        Evaluate population
        timesRun = initialTimesRun, restarts = restarts + 1
    end
end
```

represented through causal matrices which can be easily translated into Petri nets [9]. Fig. 2 depicts the Petri that models a UoL where the pedagogical objective is to learn about polymorphism in object-oriented programming.
- These individuals are then evaluated with a fitness measure that indicates how well each individual is able to reproduce the behaviour shown in the log database.
- After this evaluation process, the population evolves by selecting those individuals with a higher fitness. Then, it generates new individuals through genetic operators like crossover, which combines two individuals, and mutation, which modifies an individual.
- These steps will be repeated in a cycle until any of the termination conditions, *maxGenerations* and *maxRestarts*, are fullfilled.

In this genetic algorithm, the fitness measure ($F$) for an individual, i.e., for a learning flow, is based on its completeness ($C_f$), precision ($P_f$), and simplicity ($S_f$). When two individuals are compared to decide which of them will be selected, these three measures are considered separately:

$$\begin{aligned} F(a) > F(b) \iff & \{C_f(a) > C_f(b)\} \vee \{C_f(a) = C_f(b) \wedge P_f(a) > P_f(b)\} \\ & \vee \{C_f(a) = C_f(b) \wedge P_f(a) = P_f(b) \wedge S_f(a) > S_f(b)\} \end{aligned} \quad (1)$$

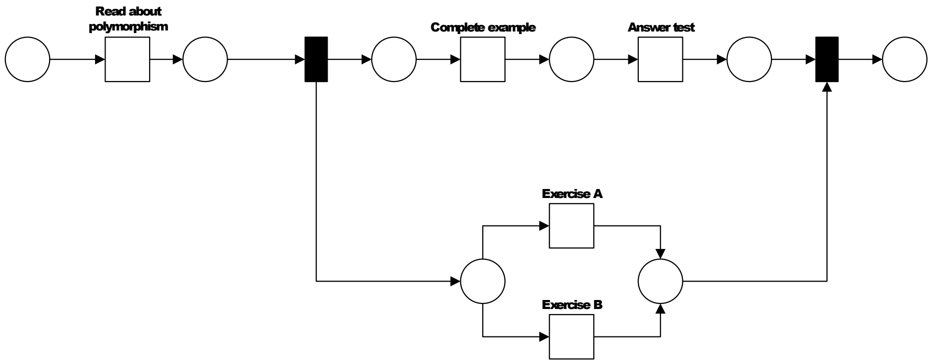



**Fig. 2.** Petri net that represents the learning flow of a UoL about polimorphism

Thus the individual with highest completeness will be selected. If the values for this measure are equal for both solutions, then the second measure to be compared is the precision. If both individuals have the same precision, the simplicity will be checked. Therefore, we are looking for the simplest learning flow from precise learning flows that are complete.

In order to check the efficiency of this algorithm for discovery learning flows, we have conducted an experiment with 14 UoLs[1] that have been undertaken in the OPENET4LD environment [12]. OPENET4LD collects all the events generated by the learners when they perform the learning activities of a UoL. Table 1 presents the results of the comparison between our genetic algorithm (GA) and two of the most used process discovery algorithms, alpha++ [14] and heuristic miner [13]. The results shows that only GA is able to discover learning paths that are complete and with high precision.

## 4 Mining Adaptive Rules from Event Logs

Using log files of VLEs can help to determine who has been active in the course, what they did, and when they did it. In this section we use these data to obtain the adaptive rules of the UoLs that determine the learning flow and the material presented to students. Unfortunately, the log files provided by VLEs do not follow a standard and therefore a generic solution cannot be formulated for such purpose. In fact, log files are seldom used mainly because it is difficult to interpret and exploit them. In most of the cases, the data aggregated are incomplete or even not logged.

Taking this into account, in this section we describe our approach based on the extract of the log file represented in Fig. 3. This example has two parts. On the one hand, the first four records represent the execution of activities and contain information about the UoL id, its name, the user that ran the activity, the level of IMS LD of this activity (play, act, role-part), the execution time,

[1] Available at `http://dspace.ou.nl`

**Table 1.** Performance of the genetic algorithm for learning flow discovery (GA)

| | | Candidas | Introduction | IMS Learning | CLFriday | Tai Chi | endolab | Learning LMS | A2 | Your Opinion | Caminatas | Trip around Spain | Debate | D2 | A8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GA | $P_f$ | 1.0 | 1.0 | 0.95 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.77 | 1.0 | 1.0 | 1.0 | 1.0 |
| | $C_f$ | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | $S_f$ | 0.34 | 0.35 | 0.29 | 0.34 | 0.34 | 0.36 | 0.36 | 0.35 | 0.37 | 0.3 | 0.35 | 0.32 | 0.31 | 0.29 |
| HM | $P_f$ | 1.0 | 1.0 | 0.96 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.73 | 1.0 | 1.0 | 1.0 | 1.0 |
| | $C_f$ | 1.0 | 1.0 | 0.75 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.2 | 1.0 | 1.0 | 1.0 | 1.0 |
| | $S_f$ | 0.34 | 0.35 | 0.29 | 0.34 | 0.34 | 0.36 | 0.36 | 0.35 | 0.37 | 0.28 | 0.35 | 0.32 | 0.31 | 0.29 |
| $\alpha^{++}$ | $P_f$ | 1.0 | 1.0 | 0.94 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.73 | 1.0 | 1.0 | 1.0 | 1.0 |
| | $C_f$ | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.3 | 1.0 | 1.0 | 1.0 | 1.0 |
| | $S_f$ | 0.34 | 0.35 | 0.31 | 0.34 | 0.34 | 0.36 | 0.36 | 0.35 | 0.37 | 0.29 | 0.35 | 0.32 | 0.31 | 0.29 |

(3, 'UOLID_201', 'SOID_0', 'GeoQuiz3', 'root', '2013-10-14 01:44:51', 'play', 'play-13193839-3ddf-41e2-a4b8-7af8dc13d059', 'play__0', 'pc_split'),
(4, 'UOLID_201', 'SOID_0', 'GeoQuiz3', 'root', '2013-10-14 01:44:52', 'act', 'act-e4901d0f-8232-4cce-a166-20e29133d279', 'act__0', 'ac_split'),
(5, 'UOLID_201', 'SOID_0', 'GeoQuiz3', 'root', '2013-10-14 01:44:52', 'act', 'act-e4901d0f-8232-4cce-a166-20e29133d279', 'act__0', 'as_split'),
(6, 'UOLID_201', 'SOID_0', 'GeoQuiz3', 'd9', '2013-10-14 01:44:52', 'role_part_role_part', 'act-e4901d0f-8232-4cce-a166-20e29133d279', 'role_part_role_part__0', 'as_start'),
...
(289, 'OPENET_RMI_SOID_0', 'UOLID_201', 'setproperty', '2013-10-14 01:59:30', 'user__2', 'd2', 'Student', 'locpers_property__12', 'Answer1', 'locpers_property', 'string', 'Venezuela'),
(290, 'OPENET_RMI_SOID_0', 'UOLID_201', 'change_property_value__0', '2013-10-14 01:59:30', 'user__2', 'd2', 'Student', 'locpers_property__1', 'Value1', 'locpers_property', 'integer', '2'),
(291, 'OPENET_RMI_SOID_0', 'UOLID_201', 'change_visibility', '2013-10-14 01:59:30', 'user__2', 'd2', 'Student', 'learning_activity__5', 'flow3', 'locpers_property', 'boolean', 'false'),
(292, 'OPENET_RMI_SOID_0', 'UOLID_201', 'setproperty', '2013-10-14 01:59:33', 'user__2', 'd2', 'Student', 'locpers_property__13', 'Answer2', 'locpers_property', 'string', 'Siria'), ...

**Fig. 3.** Example of a text-based log file

and the specific name of the activity, among other information. On the other hand, the last four records represent events on the properties used by the UoL. Specifically, each one of these records contains information about the UoL id, the operator that favor the change, the type of the property, and the new assigned value.

### 4.1 Identification of Variables and Activities

The identification of variables and activities is highly dependent on the type of events recorded in the log files. Since each VLE may record the events in a different format, the syntactic patterns used to identify these events are usually different. For example, the identified variables are highlighted in Fig. 3. In this case, a simple regular expression with the keywords "setproperty" and "change_visibility" were used to identify the variables, but a different pattern should be used to identify these same variables in an XML-based log file. However, the main issue in the identification of variables is that they are not always recorded in the log files. When this situation happens, the mechanism for learning the adaptation rules presented in this section cannot be applied.

### 4.2 Determining the Variable Values for Each Activity

In the ideal situation, each time an event is produced, *(i)* the state of the variables is saved in a log file. This means that, e.g., at the end of an activity the values of the variables of a UoL are stored in the corresponding log file. Moreover, *(ii)* in this context a value change is also considered an event and so is also recorded. However, reality is different and usually both mechanisms are not supported at the same time (e.g., in the log extract of Fig. 3 only variable changes are included).

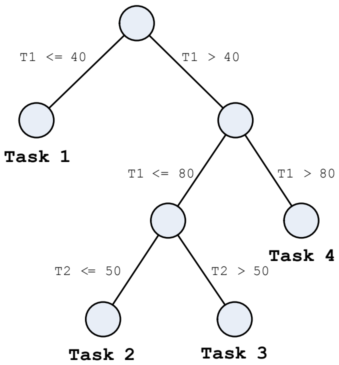
The variables values are subsequently associated to an activity, so we can determine the state of the properties before and after the activity is performed. Therefore, each time a variable value changes we must determine when and by who it was modified. In this procedure, the time of the event is crucial since it will determine the initial value of variables in the next activity.

### 4.3 Learning Rules with a Decision Tree

The identification of the adaptive rules is performed by means of J48 decision tree algorithm [10]. We selected this type of algorithms because of its simplicity, performance, and because an adaptive rule is very similar to a decision tree. As it is depicted in the left part of Fig. 4, a decision tree is a graph-like structure in which each internal node represents a test on an attribute, each branch represents the outcome of the test and each leaf node represents the class label (or decision taken after computing all attributes). A path from root to leaf represents classification rules, and in our case, an adaptive rule.

Transforming a decision tree to an adaptive rule is straightforward since these trees can easily be modelled as DNF (Disjunctive Normal Form) rules. For instance, the right part of Fig. 4 show the corresponding rules for this tree. Specifically, each branch is converted into a rule, where:

- The condition is set as the conjunction of the arc tests of a branch.
- The action of the rule is the leaf node of the branch (class/decision).
- The disjunction of all the rules has the semantics of the decision tree.




```
IF T1 <= 40
THEN visible(Task 1)

IF T1 > 40 AND T1 <= 80 AND T2 <= 50
THEN visible(Task 2)

IF T1 > 40 AND T1 <= 80 AND T2 > 50
THEN visible(Task 3)

IF T1 > 40 AND T1 > 80
THEN visible(Task 4)
```

**Fig. 4.** Transformation of a decision tree to a DNF rule base

Tree complexity has its effect on the accuracy and is usually determined by the total number of nodes, total number of leaves, depth of tree, and number of attributes used in the tree construction. The number of variables is therefore a crucial factor since too many may reduce the accuracy of the tree. Moreover, the number of data required to learn increases with the number of variables. However, adaptive rule are usually not based on many variables. Thus, our approach limits the variables considered during the learning of the decision tree, and specifically, only variables which value has changed in the execution of the activity are included.

## 5 IMS LD Reengineering

IMS LD is a well-known EML for adaptive learning which is specified in three different levels of implementation (levels A, B, and C), which determine the learning flow, the changes in the environment, and the notifications, respectively. In this paper, two learning algorithms, presented in the former sections, address the first two levels of IMS LD. Specifically, the process structure of the learning is determined by a process mining approach and the adaptation mechanisms by a decision tree. However, the results obtained by these algorithms must be transformed to adapt to IMS LD characteristics.

### 5.1 A learning Flow as a Theatre Metaphor

Particularly, the IMS LD learning flow description of IMS LD is described as a theatre metaphor where there is a number of plays, that can be interpreted as runscripts that are *concurrently* executed, being independent of each other. Each of these plays consist of a set of acts, which can be understood as a module or chapter in a course. These acts are performed in *sequence*, and in each of them the participants in the UoL carry out *in parallel* an activity or a structure of

**ALGORITHM 2.** Identification of IMS LD activity structures of type *sequence*

**Input**: $G = (T, P, F)$ is the Petri net structure, $T = A \cup C$ is the set of transitions of the net where $A$ is set of activities and $C$ the set of control transitions, $P$ is the set of places, $F \subseteq A \times P \cup P \times A$ is the set of arcs between transitions and places, $R$ is the set of roles, $U$ is the set of participants, $U_R \subseteq P \times R$ is set of user roles, $U_A \subseteq P \times A$ is the set of activities done by a user.

**Output**: SEQ is set of activity structures of type *sequence*.

1 $A_R \leftarrow \{(a, r) \mid a \in A \wedge r \in R \wedge (\forall p \mid u \in U \wedge (u, r) \in U_R \wedge (u, a) \in U_A)\}$<br>
2 SEQ $\leftarrow \emptyset$<br>
3 **for** *each* $a_i \in A$ **do**<br>
4     $S_i \leftarrow \emptyset$<br>
5     **for** *each* $a_j \in A$ **do**<br>
6         **if** $(\exists a_i, a_j, a_k, p \mid (a_i, p) \in F \wedge (p, a_j) \in F \wedge (p, a_k) \in F \rightarrow a_j = a_k)$ **then**<br>
7             **if** $\forall p_a, p_b \mid (p_a, a_i) \in P_A \wedge (p_b, a_j) \in P_A \wedge p_a \neq p_b \rightarrow (\exists r \mid r \in R \wedge (p_a, r) \in P_R \wedge (p_b, r) \in P_R)$ **then**<br>
8                 $S_i \leftarrow \emptyset$<br>
9                 **if** $\exists j \mid S_j \in SEQ \wedge a_i \in S_j$ **then**<br>
10                     $S_i \leftarrow S_j$<br>
11                 **else**<br>
12                     SEQ $\leftarrow$ SEQ $\cup S_i$<br>
13                 **end**<br>
14                 $S_i \leftarrow S_i \cup \{a_i\} \cup \{a_j\}$<br>
15             **end**<br>
16         **end**<br>
17     **end**<br>
18 **end**

activities, the which are executed in *sequence* or by *selecting* a specific number of them activities.

The problem therefore consists in identifying the different IMS LD elements in the Petri net defined by the process mining algorithm. Algorithm 2 shows one of the procedures defined for such purpose. Specifically, this algorithm is designed to identify the activity structures of type sequence in the Petri net $G = (T, P, F)$ as described in Section 3. Notice that the set $T$ is composed of two types of transitions, namely activities and control structures, where te activities represent the learning and support activities of the learning process. The identification of sequences of activities takes into account (line 6) *(i)* that there is a direct dependency, that is, an arc between an activity $a_i$ to a place $p$ and also from $p$ to a second activity $a_j$ and that here is no other dependency between $a_i$ and another activity $a_k$. Moreover, the two activities must also verify that they share the same role (line 7), otherwise they could be in different role-part structures.

The remaining algorithms used to identify the plays, acts, role-parts, or selections of activities use a very similar approach that consists in identifying the inner structural characteristics of each component. Specifically:

- *Selection of activities.* Very similar to Algorithm 2 where all the activities of the structure must share a role but also requires the activities in parallel between an *OR-split* and an *OR-join.*
- *Role-part.* The most difficult structure to identify. A role-part can be simple (just one activity with a specific role) or complex if it is composed by structures of activities, where sequences and selections can also contain subactivities and structures.
- *Act.* Role-parts must match an *AND-JOIN* workflow pattern, that is, roleparts are split in parallel by an *AND-split* control structure and their execution is synchronized by an *AND-join.*
- *Play.* Defined by a set of acts in sequence.
- *Method.* Defined by a set of plays in parallel in an *AND-JOIN* workflow pattern.

### 5.2 Knowledge-Base of Adaptive Rules

IMS LD adapts the learning flow of a student according to a set of predefined properties and rules. Properties can have different scopes, such as global or local, different targets, such as a person or people with a specific role, or even the combination of different scopes and targets, such as, e.g., global personal or local role properties. The state of a learning process is consequently defined by these variables, and therefore, depending on their values the learning flow may vary according to students' needs. Usually, the state is composed by user properties and the presentation to a student of educational material is guided by his scores, his responses to questions, or, e.g., the time needed to solve a problem. However, properties with a global scope or targeted to roles (not just to individuals) provide to IMS LD the means to adapt collaborative or groupfocused activities.

Taking this into account, the variables extracted from the event logs must be classified according to IMS LD types. The procedure detailed in Algorithm 3 returns the IMS LD scope and target of a property $p$ previously identified in the log files. Firstly, it defines the time intervals $(t, t_{+1})$ between two events that have changed the value of $p$. For each one of these intervals, the algorithm identifies the value that the property had in this instant in the log files. Specifically, we define this value for each UoL and participant (line 9). Notice that we do not have always an event from which we can obtain the value of a property in a time interval. Therefore, we store the historical values for each UoL and participant during the processing of the log files and return the last known value of the property when no event was produced in the given time interval. The property target is determined in the *if-then-else* from lines 13 to 18. In this structure we identify the cases in which the property takes the same values for each time interval. If all the participants with the same role verify this predicate then the property target is set to *role*. Notice that since this structure is in a loop all the time intervals must verify this condition to finally be a *role*-targeted property. The verification of the scope is in between lines 21 to 26. In this *if-then-else* structure we just check that the value of a property is shared between all the

**ALGORITHM 3.** Classification of a variable identified in the event logs

**Input**: $p$ is the property we want to classify, $L$ is set of logs, $M$ is the set of participants, $U$ is the set of UoLs.

**Output**: $s_p \in S$ and $t_p \in T$ are the scope and target of $p$, respectively.

**Data**: $S = \{local, global\}$ and $T = \{role, personal\}$ are the set of scopes and targets defined in IMS LD, respectively.

$t \leftarrow 0$
$V \leftarrow \emptyset$
**repeat**
  $t_{+1} \leftarrow$ next time $p$ value changed in $L$ starting from $t$
  $V \leftarrow \emptyset$
  **for** $u \in U$ **do**
    $V_u \leftarrow \emptyset$
    **for** $m \in M$ **do**
      $V_{u,m} \leftarrow \{p$ values in UoL $u$ for the user $m$ in the interval $(t, t_{+1})\}$
      $V_u \leftarrow V_u \cup \{V_{u,m}\}$
    **end**
    $V_u \leftarrow V_u \cup \{V_{u,m}\}$
    **if** $\forall m_1, m_2, r, x \mid m_1, m_2 \in M \wedge m_1 \neq m_2 \wedge V_{u,m_1}, V_{u,m_2} \in V_u \wedge r \in role(m_1) \wedge r \in role(m_2) \wedge x \in V_{u,m_1} \rightarrow x \in V_{u,m_2}$ **then**
      **if** $t_p \neq personal$ **then**
        $t_p \leftarrow role$
      **end**
    **else**
      $t_p \leftarrow personal$
    **end**
    $V \leftarrow V \cup V_u$
  **end**
  **if** $\forall u_1, u_2, m_1, m_2, x \mid u_1, u_2 \in U \wedge u_1 \neq u_2 \wedge m_1, m_2 \in M \wedge m_1 \neq m_2 \wedge V_{u_1,m_1} \in V_{u_1} \wedge V_{u_2,m_2} \in V_{u_2} \wedge V_{u_1}, V_{u_2} \in V \wedge x \in V_{u_1,m_1} \rightarrow x \in V_{u_2,m_2}$ **then**
    **if** $s_p \neq personal$ **then**
      $s_p \leftarrow global$
    **end**
  **else**
    $s_p \leftarrow local$
  **end**
  $t \leftarrow t_{+1}$
**until** $t \neq t_{+1}$

UoLs in the time interval. If it is, the scope of the property is set to *global*, and to *local* if it is not.

The second pillar of IMS LD adaptation mechanism is a set of rules. Rules are structured as a conditional *if-then-else* control flow and therefore evaluate the *then* part when the *if* condition is verified and the *else* part when it is not. The operators used in the *if*, *then*, and *else* blocks are defined in the IMS LD

Level B of the information model [3] and include classic logical, arithmetic, and comparison operators, and others more specific that support operating on the user model and the process structure (i.e., methods, acts, role-parts, and activities). Contrary to programming languages, the semantics of the *then* part is limited and so can only contain actions, which will show or hide objects, change a property value, or notify a role. The *else* part is more standard, and allows concatenating *if-then-else* control structures, in addition to actions.

Notice that IMS LD rules structure is very similar to those described in Section 4. Two additional remarks about the transformation of the decision tree. Firstly, there is always a viable transformation between the obtained decision tree and the IMS LD rules since *(i)* the operators used by the decision tree are supported by IMS LD and *(ii)* we just learn two types of actions for changing *(i)* the visibility of an object and *(ii)* the value of a property. Secondly, the rules we generate have the structure of an *if-then* control flow, that is, we never have an *else* block. This last feature does not limit the semantics of our rule model since each *else* is represented as an *if-then* rule but with its specific condition.

Finally, we must also mention that in IMS LD rules are evaluated at the beginning of a session and each time a property is changed. Furthermore, rules are also evaluated at the completion of methods, plays, acts, and activities. This is reasonable since these control structures coordinate the participants of the UoL and thus that or teachers change the value of their properties in the activities, methods, plays, and acts define a group behaviour and thus changes must affect all the members of the group. However, this behaviour is very dependent of IMS LD and even more complicated to extract from the events of log files. In fact, usually this kind of adaptation is represented by the process mining as a simple activity that changes the state of the properties. Consequently, the completion of upper hierarchical structures, such as plays and acts, will not change the state of learning. However, this is not a big issue since our solution defines specific dummy activities to such a change.

## 6 Conclusions and Future Work

In this paper we have proposed a global approach that facilitates the reuse of UoL defined in legacy systems or VLEs. Our solution has been implemented to support the mining of event log files (i) to obtain the learning flow, formalized as a Petri net, and performed by students and instructors, and (ii) to identify the adaptation rules, represented as decision trees, used to tailor the learning flow to each student. From these two results, we also describe how to translate these models into an IMS LD UoL.

Finally, as future work we plan to extend the number of operators used by the decision tree algorithm to so give support to the complete grammar or IMS LD.

## References


1. van der Aalst, W.M.P.: Process Mining: Discovery, Conformance and Enhancement of Business Processes, 1st edn. Springer Publishing Company, Incorporated (2011)
2. Bergenthum, R., Desel, J., Harrer, A., Mauser, S.: Modeling and Mining of Learnflows. In: Jensen, K., Donatelli, S., Kleijn, J. (eds.) ToPNoC V. LNCS, vol. 6900, pp. 22–50. Springer, Heidelberg (2012), http://dl.acm.org/citation.cfm?id=2231056.2231058
3. Global Learning Consortium, I.: IMS Learning Design Information Model, Version 1.0 Final Specification (March 2003), http://www.imsglobal.org/learningdesign/index.html
4. Dalziel, J.: Implementing learning design. the learning activity management system (lams). In: Proceedings of the 20th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE), pp. 51–58. IADIS Press, Adelaide (2003)
5. 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 (First 2012)
6. Hernández-leo, D., Villasclaras-Fernández, E.D., Asensio-Pérez, J.I., Dimitriadis, Y., Jorrín-Abellán, I.M., Ruiz-Requies, I., Rubia-Avi, B.: Collage: A collaborative learning design editor based on patterns. Educational Technology & Society 1(9), 58–71 (2006)
7. Karampiperis, P., Sampson, D.: A flexible authoring tool supporting adaptive learning activities. In: Proc. of IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2004), pp. 51–58. IADIS Press, Lisbon (2004)
8. Martinez-Ortiz, I., Sierra, J.L., Fernandez-Manjon, B.: Authoring and reengineering of ims learning design units of learning. IEEE Trans. Learn. Technol. 2(3), 189–202 (2009), http://dx.doi.org/10.1109/TLT.2009.14
9. Murata, T.: Petri nets: Properties, analysis and applications. Proceedings of the IEEE 77(4), 541–580 (1989)
10. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco (1993)
11. Reimann, P., Yacef, K.: Using Process Mining for Understanding Learning. In: Handbook of Design in Educational Technology, pp. 472–481. Routledge, New York (2013)
12. Vidal, J.C., Lama, M., BugaríN, A.: Petri net-based engine for adaptive learning. Expert Syst. Appl. 39(17), 12799–12813 (2012), http://dx.doi.org/10.1016/j.eswa.2012.05.013
13. Weijters, A., van der Aalst, W., Alves de Medeiros, A.: Process mining with the heuristics miner-algorithm. BETA Working Paper Series WP 166. Eindhoven University of Technology (2006)
14. Wen, L., Aalst, W.M., Wang, J., Sun, J.: Mining process models with non-free-choice constructs. Data Min. Knowl. Discov. 15(2), 145–180 (2007), http://dx.doi.org/10.1007/s10618-007-0065-y

# Moodle4SPOC: A Resource-Intensive Blended Learning Course

Sabrina Ziebarth and H. Ulrich Hoppe

University Duisburg-Essen,
Department of Computer Science and Applied Cognitive Science
Lotharstr. 63/65, 69121 47048 Duisburg, Germany
{ziebarth,hoppe}@collide.info
http://www.collide.info/

**Abstract.** The paper reports on the transformation of a master level university course (lecture style) by supplementing it with a "small private online course" (SPOC). The online course included micro learning elements and activities that are often used in the context of MOOCs like lecture videos or auto-grading tests, but also provided a platform for publishing content that was generated by the students in group work. Technically, the online course was conducted in a Moodle LMS that was extended by a self-developed video-plugin. The experience was evaluated using two questionnaires and the results were discussed with the students to gain further insights. The evaluation results are critically discussed to identify success factors and pitfalls for resource-intensive blended learning scenarios.

**Keywords:** blended learning, small private online courses, lecture support, lecture videos, group learning, peer reviewing.

## 1 Introduction

The Department of Computer Science and Applied Cognitive Science at the University of Duisburg-Essen (Germany) offers three different study programmes: "Applied Computer Science", "Interactive Media and Cognition", and "Computer Engineering" (all Bachelor and Master levels) with a total of about 1500 students. The Computer Engineering programme is mainly taught in English, whereas the standard of the other programmes is German. In this context, GILLS ("**G**estaltung **i**nteraktiver **L**ehr- / **L**ern**s**ysteme") is a Master level course that focuses on user/task modeling and models of interactive and collaborative learning environments. It targets students from the three Master programmes and is also eligible by higher education teacher students of computer science, a programme offered by another faculty. The course is taught once per year with about forty participants.

The course consists of a weekly three hour lecture accompanied by an one hour exercise with weekly assignments. Although the GILLS lecture is given in German, all presentations (interactively annotated slides) and additional reading materials are provided in English to support the international Computer

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 359–372, 2014.
© Springer International Publishing Switzerland 2014

Engineering students. All course materials, assignments, announcements as well as a Q&A forum are offered through a Moodle portal. Originally, the three hours of lecture as well as the one hour of exercise were conducted consecutively as a four hour block with two breaks (one during the lecture and one between lecture and exercise).

Probably due to the given schedule of the course, attendance in the exercise was typically lower than in the lecture and the willingness for active participation (i.e. presentation of results on the part of the students) was also low. This led us to the idea of "virtualizing" the exercise, i.e., to replace the presence based activity by a combination of individual and group assignments with reporting back into the Moodle platform. The assignments would be partly based on reading materials (summary, analysis) and partly be oriented towards constructive problem solving and modeling. The individual activities would also be supported by quizzes (using Moodles quiz mechanisms). In this new concept, also the role of the lecture was redefined in the sense of providing orientation knowledge ("advance organizer") and core definitions with a limited number of characteristic examples, leaving specific extensions to the virtual exercise activities. These lecture elements were provided as smaller thematic video clips, available on the platform. Additionally, the plenary face-to-face setting of the lecture was used as a forum to take up and discuss results from the virtual exercise activities. Thus, the focus of the classroom experience shifted from teacher centered lecturing to student centered interactions.

This course redesign led to a blended scenario with cooperative elements and content production of the part of the students. The use of quizzes and videos corresponds to instructional formats typical of MOOCs ("massive open online courses"). Given that our course is neither open to a broader public beyond the three mentioned master programmes nor "massive" in terms of participation it could rather be categorised as a SPOC ("small private online course") in the sense of Fox [1]. According to Fox, SPOCs intend to "increase instructor leverage, student throughput, student mastery, and student engagement". This is achieved partly by adopting learning elements and activities also used in xMOOCs such as lecture videos, reading materials, weekly exercises and forums for questions and discussions. Due to potentially unlimited participation, in MOOCS these exercises consist of tasks that can be automatically graded such as online quizzes or tests; also peer reviews are used for feedback and grading (cf. [2]). Regarding collaborative task assignments, our approach also shares elements with connectivist MOOCs (or cMOOCs) as introduced by Siemens and Downes (cf. [3]). Such cMOOCs focus on the production and sharing of knowledge artifacts like blog or wiki articles by and between learners: "In connectivism, the starting point for learning occurs when knowledge is actuated through the process of a learner connecting to and feeding information into a learning community" [4].

Meanwhile there are numerous examples of extending and enhancing traditional learning scenarios with online and interactive elements, usually subsumed under the notion of "blended learning". Graham [5] distinguishes two types of blended learning scenarios: "enhancing blends" denotes the usage of online tools and resources as an add-on to existing models, whereas "transformative blends" would be those in which the pedagogical model is transformed together with the introduction of the technologies. Such transformations can be related to role models for students and teachers (as in the "flipped classroom" approach - cf. [6]), the overall organization and workflow of the course or new types of core activities in the course.

Our approach is clearly transformative in the sense of Graham. From a research point of view, we conceive our course as an experiment through which we want to investigate the effectiveness of the newly introduced learning resources and activities. A recent report to the U.S. Department of Education [7] summarises empirical evidence on the effects of online learning practices in the form of a meta-analysis. Although this report supports the overall effectivenes of online learning, it questions a positive effect from online quizzes and videos. If generalizable, this would be crucial for a variety of formats of online instruction, including MOOCs and other types of web-supported courses. In our study we particularly focus on videos, quizzes/games, and collaborative content production. In addition to the observation of actual usage we are particularly interested in the acceptance and valuation of these resources by the students. E.g., would teacher-created materials (videos, quizzes) be preferred to student-created content (wiki articles) in the preparation of the exams? To investigate these questions we rely both on the analysis of log data and on student responses to questionaires.

## 2 Experimental Setting: Learning Plattform, Materials, and Procedure

The modified version of the GILLS course was first offered and conducted in the summer semester 2013. The virtualized exercise was conducted in our Moodle 2[1] LMS mainly based on existing built-in functionality. We additionally used a self-developed video-plugin (see figure 1), which allows for setting bookmarks for certain scenes in the video and for associating tags that can serve as navigation "bridges" between videos.

The usage of Moodle for SPOCs provides the following advantages:

- teachers as well as students are familiar with the learning environment
- it is open source, thus departments can set-up their own instances
- there are many available plugins for extending and adapting Moodle
- if needed, own plugins can be created to influence learning processes

Lecture videos were audio recorded and grabbed from screen (with Camtasia[2]), i.e. not showing the lecturer/presenter as part of the video. Videos were recorded and sometimes cut in such a way as to form smaller learning units centered around specific topics. For lecture support a kind of digital blackboard

[1] http://moodle.com/
[2] http://www.techsmith.com

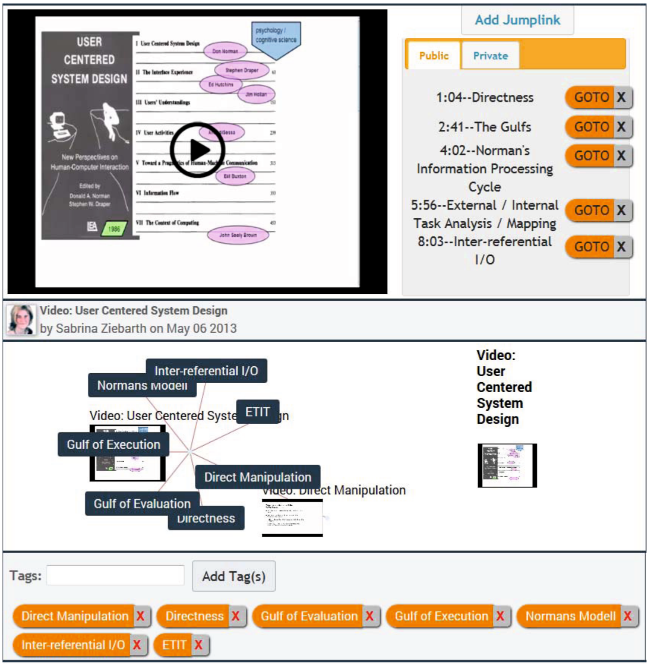



**Fig. 1.** Video plugin in Moodle.

was used consisting of a convertible tablet computer, the FreeStyler[3] whiteboard application and a beamer. Thus, slides could be interactively edited regarding dialogue with the students, but also be delivered to the students.

The Moodle course contains one forum for announcements and one for questions and discussions of the participants. The students were asked to post questions relevant to the whole course in the forum rather than to write emails to the supervisors. Additionally, the forum was used to support group exercises. Furthermore, a wiki was introduced to build a glossary for repetition and exam preparation. We did not use the Moodle glossary since it does not support co-operative editing and the corresponding articles to relevant topics of the lecture were created collaboratively by groups of students. A typical exercise regarding a wiki article consisted of one week of creating the initial version in a small group, one week of individually reading the articles of other groups as well as writing peer feedback and one week of revising the article in the original small group. Each group handled another topic enabling students to specialize regarding their interests. Thus, groups were formed by topic using Moodles choice activity, which was set up to ensure groups of similar size. Furthermore, students could see the choices of their peers to enable group formation and communication. In addition

[3] `http://projects.collide.info/projects/freestyler-release`

to the peer reviews, additional feedback was given by the supervisors for quality assurance. Overall, the online exercises contained (per student):

- Creating and revising four wiki articles in small groups
- Two task regarding constructive problem solving and modeling (partially using external tools) in small groups
- Writing peer reviews (as wiki comments) regarding the group results above (individual task)
- Creating and extending a concept map regarding the topics of the first chapter in small groups using external tools like bubbl.us[4]
- Three quizzes regarding the three main topics for self-testing, which could be completed several times
- Creating own questions and sample solutions in a wiki with regard to the lecture topics
- Playing the serious game Matchballs [8] for training relations between important concepts of the lecture (see figure 2)

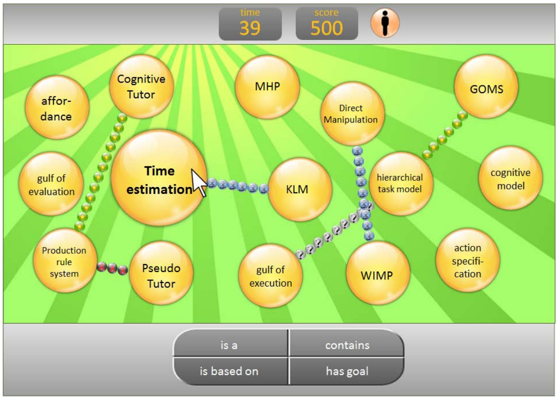



**Fig. 2.** Interface of the serious game *Matchballs* (translated to English)

Matchballs [8] (see figure 2) is an ontology based learning game, in which learners have to create relations between given concepts (represented as balls). Thus, it is similar to a quiz, that trains factual knowledge. The game can be played by two players or by one player and a bot that simulates the second player based on the knowledge stored in the ontology. Players get points for relations they agree on, i.e. if both create the same relation between two concepts. They may see the relations of their teammates, but not the relation types. The relations still to be matched by the other player are visualized by hour glasses. The relations proposed by the teammate are shown as lines of question marks. If the teammates agree on a relation, this relation is displayed by stars, if they

[4] `https://bubbl.us/`

disagree, it is marked by red Xs. In addition to training the students, Matchballs harvest their understanding of the domain to find potential new relations that are missing in the ontology as well as potential misconceptions. Thus, it is also a game with a purpose (GWAP) (c.f. [9]).

To motivate students for participating in the online exercises, in addition to having a more interactive and student-centered course bonus points were introduced. Furthermore, half of the oral exam would be about topics the students dealt with by themselves in the exercises (e.g. the created wiki articles).

## 3 Evaluation

The initial number of participants was a bit higher than usual and drop-out rates were very low. The first lecture was visited by 48 students instead of around 40 in the past years. Of the 44 students doing the first exercise 40 regularly did their exercises till the end of the course, one dropped out after the first exercise, two in the middle of the course and one after two thirds. Each student achieved 80 % of the bonus points on average, not considering the four drop-outs the students had an average of 85 % of the points. 37 of the 40 students participating in the exercises registered for the oral exam, the other three decided to do it later. 36 exams took place, one was cancelled, the average grade was 1.48 with bonus points and 1.99 without (best possible grade is 1.0 and worst passed grade 4.0).

### 3.1 Data Analysis

The Moodle course was frequently visited: There were 24 distinct students per day and 45 per week on average (see also figure 3). The highest amount of distinct users per day was on lecture days and the lowest on weekends. The deadline for and assignment of new exercises was typically on lecture days, which may explain the high number of accesses on these days.

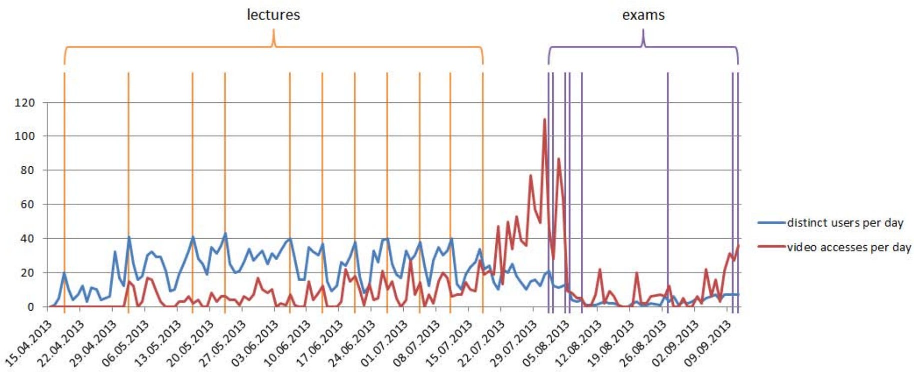



**Fig. 3.** Moodle usage

There were overall 26 lecture videos which could only be accessed in the Moodle platform. The number of video accesses increased highly towards the beginning of the exam period (see figure 2), thus we assume that they were used for exam preparation. Each exam relevant video was accessed by 37 distinct students on average; the last two non-exam relevant videos were accessed by 15 distinct students on average.

The forum was used for questions regarding the course (26 threads) as well as the coordination of group exercises (22 threads). The course related questions mainly referred to organization (e.g. late registration for groups) as well as Moodle usage and problems (e.g. how to add media to wiki articles). Most questions that could be answered by students were answered by them, which is an indicator for an active course community.

Regarding the wiki, 28 articles were created by student groups. These articles were mainly very detailed and of high quality, and thus could be used for exam preparations by the whole course. We ordered them alphabetically using the glossary analogy, but one student also autonomously created an additional ordering of the articles based on the course structure (chapters and sections) for exam preparation.

In general, peer reviews were very good and based on the given questions. By combining all reviews regarding one article most problems were identified (often also relevant ones the supervisors did not find). One problem with the reviews was based on the usage of the wiki comment function. Since everybody could see every comment, the last reviewers often found nothing new resulting in comments like "I go along with the other reviews, all relevant points already have been stated".

Not only the tasks with bonus points were solved, but also the voluntary ones: 30 students did the first quiz, 25 the second one as well as the third one. The Matchballs game was played by 19 students between 1 and 34 times (ca. 8 games per user on average, standard deviation of 7.85). Most games were played singleplayer, but there were also two multiplayer games involving three different players. Overall 454 distinct relations were played, 99 of them already were in the ontology and 355 not. Most of the new relations were only played by one or two players. Only new relation that have been played by at least a minimum number of different players are considered for finding new relations at all. This ensures that a new relation has a minimum support. In this experiment we chose a minimum support of four and found that all new relations that fullfilled this threshold condition were true and could be integrated in the ontology. Thus, regarding the gwap concept, we could indentify new knowledge for our system. The question, if the students learned something by playing the game, is not so easiliy answered. We provided a small questionaire that was only completed by eight students, thus the number of students' assessements is very low. 75 % of the answers indicate, that Matchballs is appropriate for recapitulation of learning contents and 50 % of the student partially agree that they learned new relationships regarding the learning content. The exam results show that the students using the game were slightly better (average grade of 1.2) than the ones not using the game (average grade of 1.5). Within the group of students, who played the game, the ones who played it more often also tend to have better grades than the ones, who only played it only one to three times. Of course, highly motivated/hard-working students tend to use additional opportunities for learning more often and tend to get better grades, which might explain the above average grade results.

### 3.2 Resource Access Patterns

Based on the Moodle logs, we have applied network analysis techniques to detect the grouping of students around resources. The primary relation in this analysis associated students to resources, thus forming an actor-artifact network. The "biclique communities method" has allowed us to detect cohesive subcommunities in the original actor-artifact network [10], i.e. particularly dense groupings of students around certain resources. Based on additional heuristics we have traced the evolution of these subcommunities over the whole lecture and exam period. During the time of the lecture, the resource access patterns reflected mainly the groups induced by task assignments. We also found one grouping of students around a set videos, which could not be explained by the assignments.

The analysis of the exam preparation phase (after the end of the lecture period) was less predictable, because this phase was self-organized by the students. In this phase, a core set of learning materials (lecture slides regarding the last lecture topics relevant for the exam and quizzes for self-testing) were used by the majority of the students. Most of the oral exams took place two to three weeks after the end of the lecture period, but due to organizational constraints the exams of one study programme (Computer Engineering - CE) were held six weeks later. This resulted in a gap in the resource access figures after the first exam period (see also figure 3). Students of the second exam phase showed a different resource access patterns for exam preparation as compared to the other groups. The resource access figures for the CE group indicate that they started their exam preparation later and overall appeared to invest less time. Also, they began their preparation by focusing on lecture videos and/or wiki articles as compared to the lecture slides preferred by the other groups. Only at the end of their preparation phase, the CE group shifted their focus to the materials used by the others. One interpretation of these behavoral differences is that the first group started their exam preparation in the lecture period, i.e. these students only had to recapitulate the last topics, for which there were no wiki articles. The second group had to recapitulate all topics. The time invested for exam preparation might also explain, why students from the first group tended to get better grades (1.3 on average) than students from the second group (2.1 on average).

Figure 4 shows the inter-relation of students in the course based on their grouping in terms of shared or overlapping grouping resource usage. The edge width indicates how often two students were in the same group (the thicker the edge, the stronger the relation). The node size represents the centrality of the students. The network diagrams suggest a tendency for students to work

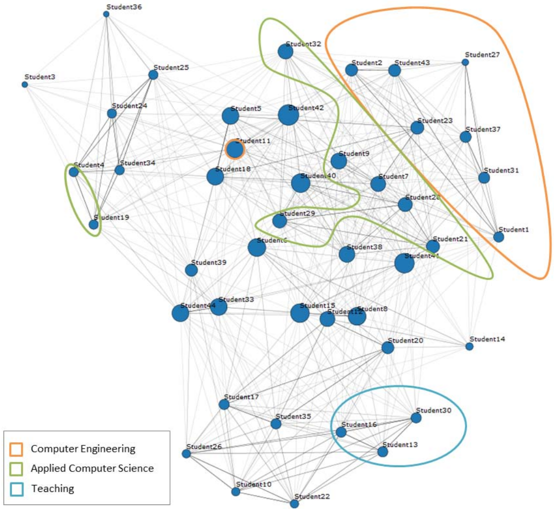



**Fig. 4.** Course community based on group work

together with peers from the same study programme. Students of Computer Engineering reported that they had problems finding groups beyond the peers from their programme. This is reflected by the network, in which students of Computer Engineering (orange frame) are at the periphery of the network and tend to have strong ties among each other. Students of Applied Computer Science (green frame) seem to have a mediating role between the students of Computer Engineering and Interactive Media and Cognition (no frame). The network has a density of 0.225, meaning that students worked together with 22.5(around 10) of their peers on average.

### 3.3 Student Feedback

To gain further insights into the specifics of the course an online questionnaire was conducted that was answered by 24 students. The students had to rate their agreement regarding given statements based on a scale ranging from full disagreement (1) to full agreement (5). The results are displayed in figure 5 and 6.

The course was perceived as interactive and the majority of students stated that they engaged more than in other courses in the lecture period. Despite their high engagement, students seemed to be unsure regarding their preparedness for the exam. Lecture videos as well as additional videos were stated to be beneficial for learning and lecture videos as well as slides were considered as main resource for exam preparation. The usefulness of the forum for solving individual as well as group tasks was controversial. Given quizzes as well as peer generated ones (q/a wiki) were mainly perceived as useful for learning and students planned to use them for exam preparation (75 % given quizzes, 58 % peer generated quizzes).

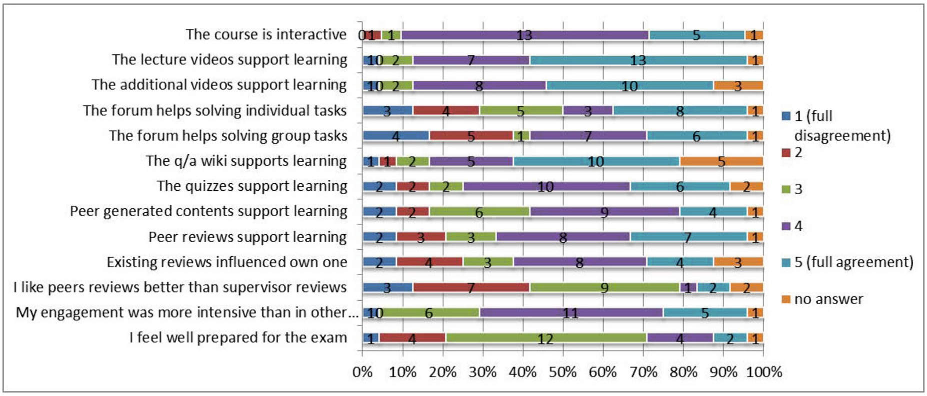



**Fig. 5.** Feedback regarding provided tools and tasks

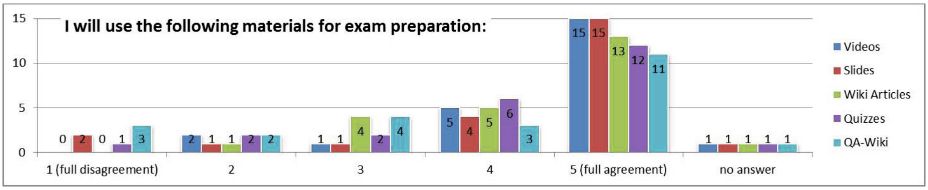



**Fig. 6.** Planned material usage

Similarly, peer generated contents (wiki articles) were described to be useful for learning by the majority of students and considered as resource for exam preparation. While students stated the contents of the peer reviews to support learning, they did not like to be reviewed by peers more than by supervisors. Half of the students declared that their reviews were at least partially influenced by the already existing reviews of their peers.

While teacher-generated contents (videos and slides) score slightly higher as materials for exam preparation than student-generated contents (wiki articles), there are no conclusive differences between them. This indicates a high confidence in peer-generated content. To ensure high quality wiki-articles, they were controlled by teaching assistants. Furthermore, students were answered in affirmative that the wiki articles could be used for learning. We think that the quality assurance of the student-generated contents was very important for their acceptance.

The course was additionally evaluated with the universitys standard course evaluation questionnaire (paper based), which was completed by 19 students in the lecture. The students stated that they visited the lecture regularly (79 % almost always and 21 % often) and also often participated in the exercise (79 % almost always, 10.5 % often, 10.5 % no answer). The course was experienced as challenging: the lecture was perceived as appropriate by 47.4 %, as difficult by 47.4 % and as too difficult by 5.3 %; the exercise as appropriate by 42.1 %

and as difficult by 52.6 % (one student did not answer). All in all, the course scored below average (2.58 on a scale from 1 (best) to 5 (worst) with an average of about 2.0 in the engineering faculty). This appears to be contradictory to the very low drop-out rate between initial participation and examination, given that the course was not obligatory and alternative courses with different content were available. One issue that was raised in individual statements was that some students had expected a more practical introduction to designing and building standard e-learning environments instead of focusing on (partly formal) models and analysis techniques.

Another reason for not giving a better overall score was an expression of complaint about the workload required for the exercises. The reported efforts were distributed as follows: 5.3 % less than 1 hour, 21.1 % 2-3 hours, 26.3 % 3-4 hours, 36.8 % 4-5 hours, 5.3 % 5-6 hours and 5.3 % more than 6 hours. Evidently, many students considered this as not acceptable. Free input fields of both questionnaires as well as discussions with the students showed, that the main effort was perceived in the coordination of the group work. Groups were stated to be too large. The group size for the wiki articles was seven students on average (standard deviation ca. 2) and for the other group tasks (concept maps, problem solving/modeling) four students (standard deviation ca. 1). To reduce coordination effort the maximal group size asked for was three to four students. Furthermore, students did not like the new formation of groups based on topics, since they considered the alternating groups to increase their coordination effort and conflict with long term division of work. Another problem was that bonus points were gained for group work, thus there was a dependence on peers. An additional reason for the bad grade was the gulf of students expectation regarding the course and the actual content. In spite of the critique there was also positive feedback regarding the course concept (interactive lecture, wiki articles and peer reviews, group work).

## 4 Discussion

The engagement of the students in the virtualized exercise with MOOC elements was much higher than in the previous presence based one. Despite of the provision of videos the lecture was well visited and students were ready and willing to participate actively. The videos seem to have been used especially for exam preparation. There is typically a number of students who visit many courses in the first weeks of the lecture period to decide which ones to take, thus the number of students is higher in the first weeks. In GILLS, a high proportion of students appear to have "stuck" with the course. Although there was much critique regarding the effort, the drop-out rate was very low and exercises were done regularly with high quality outcomes. Asking the students why they did not drop-out of the elective course showed that after having initially invested effort (for the first exercise) they were not willing to "lose their investment". Thus, bonus points seem not only to be highly motivating, but also kind of binding for students. Nevertheless, students did not only do the exercises with bonus points, but also the ones without (quizzes and serious game) that were recommended for repetition and exam preparation. Self-tests (quizzes) were received as beneficial for learning, which is similar in MOOCs (cf. [11]).

The course was not only a high effort for the students, but also for the supervisors, since the exercise activities were completely redesigned. Basic reading material as well as central questions for 28 wiki articles were provided and the resulting wiki articles had to be quality-ensured, each addressing another topic and going in much detail. Each group or individual task (despite quizzes and the serious game) was manually corrected and graded. Group work is not only considered beneficial for learning (knowledge sharing), but also reduces the number of submissions, which have to be corrected. Even so, groups have to be really small to focus on learning and not mostly on group coordination. Thus, it seems reasonable to give smaller tasks to smaller groups than bigger topics to larger groups, and demand cross-references of the articles to support the linkage of the user-generated knowledge sources. Since peer reviews were very good, they seem to be a good alternative to supervisors feedback. However, while students considered the peer reviews as helpful, they did not really trust the assessment of their peers and very much opposed the idea of being graded even only for bonus points by their peers. The student-generated wiki articles seem to be accepted as learning sources (see figure 6), but students were initially uncertain and quality assurance of the supervisors was required. The creation of good, unambiguous quizzes containing multiple choice as well as cloze elements is a lot of work. Enabling students to generate their own quizzes, not only produces a much larger pool of questions, but also addresses the idea of "learning through teaching" [12] or "constructionism" [13]. However, also for these quizzes, non-peer quality control is requested for acceptance.

Moodle proved to be only partially adequate for our goals, since there is no real support for peer reviewing, thus we used the comment function of the wiki. Since all comments can be seen by all students, the reviews were biased and students were confronted with finding new problems that their peers had not already stated. Furthermore, due to the high number of features, Moodle also has usability issues, resulting in problems when students were asked to use it for the generation of contents (e.g. wiki articles). On the other hand, Moodle is often used for online/blended learning, offers many relevant and useful built-in features, and the participating students were already familiar with it.

## 5 Conclusions and Open Issues

The students perceived the transformative nature of the blended learning course in judging it as interactive (86 % agreement, see figure 5) and highly engaging (66 %). Regarding our research questions we found that teacher-created materials as well as student-created content were generally accepted and regarded as beneficial for learning by the students. There is a slight preference for teacher generated content (83 % of the students planned to use the videos, 79 % the slides and 75 % the quizzes), but the majority (75 %) also considered the student-generated wiki articles and 58 % the student-generated quizzes (QA-wiki) as sufficient learning resources. 83 % consider the videos to support learning, 66 %

the quizzes and after all 54 % student generated contents. Data analysis showed that the videos were intensely used for exam preparation. Thus despite having no empirical proof on the learning gain, we consider online quizzes and videos as beneficial for learning due to their acceptance and valuation by the students.

Based on the evaluation results, we feel encouraged to continue and further develop the introduced approach for this and other lectures in a mainly presence based campus setting. In spite of less positive ratings regarding the efforts required for group work, we would not scale down this requirement. Yet, overheads in group formation and group coordination have to be reduced and group size of more than seven members should be avoided.

A very important element in our course was the plenary presentation and discussion of the results of group assignments in the lecture setting. Some of the presentations paraphrased and partly extended concepts previously addressed in the lecture and elaborated in the wiki. Here, the lecturer typically assumed the role of questioning certain assumptions to detect misunderstandings or intellectual simplifications ("beyond the obvious"), often stimulating intensive discussions. Another approach was to use group assignments for creating experts in different subtopics and combine them in discussion groups in a Jigsaw design (cf. [14]). For example, after writing wiki articles on certain aspects of media and learning theory in small groups, new groups containing experts in all of the considered aspects were formed in the classroom to discuss "ideal" learning systems for certain scenarios based on media and learning theories. The results were again presented and discussed in plenary.

Particularly related to modeling activities, we have also tried to invoke active "performances". One of the highlights was, e.g., the enactment of production system based cognitive task models in combination with a device model (of a vending machine) implemented as a Petri net. Here, the rule-based cognitive model was "enacted" in the form of a role play (a player representing a production rule).

If we try to extrapolate our experience to real MOOCs, most of our findings regarding the usage of videos, group work and peer reviewing appear to be directly relevant. A harder challenge would be the implementation of an equivalent to plenary presentations and performances. The kinds of desirable interactions that we have seen here depend on partly spontaneous contributions and exchange in a synchronous, real time setting. In remote scenarios it might be possible to replace this scenario by real time sessions in which presentations or performances are substituted by videos previously prepared by the learners. These videos would be watched and commented together with strong interventions and stimuli provided by a moderator.

To support the better understanding of complex, theoretical concepts, two virtual labs[5] have been developed that will be used in the next instance of the GILLS lecture in summer term 2014. Virtual labs can promote conceptual understanding by allowing students to "explore unobservable phenomena; link observable and unobservable phenomena; point out salient information; enable

[5] http://collidelabs.collide.info/

learners to conduct multiple experiments in a short amount of time; and provide online, adaptive guidance" [15]. The goal of the GILLS labs is to provide hands-on experience regarding "cognitive modeling based on production rule systems" and "knowledge diagnosis based on Bayes-nets" through interactive simulations. The simulations allow students to explore phenomena (e.g. cognitive processes) that cannot be observed in real life. To prevent installation effort and ease maintenance, the labs were implemented as web applications. Virtual labs do not only provide an supplement for SPOCs, but are also interesting for MOOCs, since they enable self-regulated, active exploration of theoretical concepts independently of time and space.

## References

1. Fox, A.: From moocs to spocs. Communications of the ACM 56(12), 38–40 (2013)
2. Belanger, Y., Thornton, J.: Bioelectricity: A quantitative approach - duke university's first mooc (2013), `http://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/6216/Duke_Bioelectricity_MOOC_Fall2012.pdf`
3. Fini, A.: The technological dimension of a massive open online course: The case of the cck08 course tools. International Review of Research in Open and Distance Learning 10(5) (2009)
4. Kop, R., Hill, A.: Connectivism: Learning theory of the future or vestige of the past? International Review of Research in Open and Distance Learning 9(3) (2008)
5. Graham, C.R.: Blended learning systems: definition, current trends, and future directions. In: The Handbook of Blended Learning: Global Perspectives, Local Designs. John Wiley & Sons (2006)
6. Lage, M., Platt, G., Treglia, M.: Inverting the classroom: A gateway to creating an inclusive learning environment. Journal of Economic Education 31, 30–43 (2000)
7. Means, B., Toyama, Y., Murphy, R., Bakia, M., Jones, K.: Evaluation of evidence-based practices in online learning: A meta-analysis and review of online-learning studies. Technical report. US Department of Education, Washington, DC (2009), `http://eric.ed.gov/?id=ED505824`
8. Ziebarth, S., Malzahn, N., Hoppe, H.U.: Matchballs – A multi-agent-system for ontology-based collaborative learning games. In: Herskovic, V., Hoppe, H.U., Jansen, M., Ziegler, J. (eds.) CRIWG 2012. LNCS, vol. 7493, pp. 208–222. Springer, Heidelberg (2012)
9. von Ahn, L.: Games with a purpose. Computer 39, 96–98 (2006)
10. Hecking, T., Ziebarth, S., Hoppe, H.U.: Analysis of dynamic resource access patterns in a blended learning course, Indianapolis, Indiana, USA, pp. 173–182. ACM Press, New York (2014)
11. Grünewald, F., Meinel, C., Totschnig, M., Willems, C.: Designing mOOCs for the support of multiple learning styles. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 371–382. Springer, Heidelberg (2013)
12. Gartner, A., Conway Kohler, M., Riessman, F.: Children teach children. Learning by teaching. Harper & Row, New York (1971)
13. Papert, S., Harel, I.: Constructionism. Ablex Publishing Corporation, Norwood (1991)
14. Aronson, E., Patnoe, S.: The Jigsaw Classroom: Building Cooperation in the Classroom, 2nd edn. Longman, New York (1997)
15. de Jong, T., Linn, M.C., Zacharia, Z.C.: Physical and virtual laboratories in science and engineering education. Science 340, 305–308 (2013)

# Requirements for Supporting School Field Trips with Learning Tools

Madiha Shafaat Ahmad, Nguyen Thinh Le, and Niels Pinkwart

Department of Computer Science, Humboldt-Universität zu Berlin, Berlin, Germany
{madiha.shafaat.ahmad,nguyen-thinh.le,pinkwart}@hu-berlin.de

**Abstract.** This paper attempts to identify the key requirements for learning tools that facilitate learning processes both in formal classroom settings and on outdoor field trips. For that purpose, a qualitative study has been conducted that consisted of interviewing ten teachers from two different states of Germany. The study showed that field trips are a combination of formal and informal learning and could be better facilitated by utilizing heterogonous learning technologies based on particular stages of a field trip. This study suggests a list of requirements for learning tools that facilitate field learning in multiple domains.

**Keywords:** Mobile learning, Collaborative learning, Learning technologies.

## 1 Introduction

Fields trips represent an established learning and teaching mode in education. One of the main focuses of mobile learning is to facilitate field trips and informal learning. Although numerous examples [1, 2, 3, 4] of mobile learning have addressed the requirements of field learning, most of them have some limitations. For instance, many of proposed tools were domain specific and deeply contingent on specific learning scenarios, e.g., museum visits. Furthermore, they often strongly depended on a particular learning theory and presented lack of flexibility for integrating or switching between various learning theories (or teaching approaches). As field trips are often an integration of informal learning elements into formal learning, combining mobile technology and in-class learning technologies (e.g., interactive whiteboards and tables) would certainly be beneficial. Integrating the heterogeneous landscape of these technologies seamlessly into learning processes would present an excellent opportunity to facilitate both learners and teachers. Thus, we tried to reexamine the requirements for integrating various learning technologies in field learning. For that purpose, we conducted a qualitative study that was comprised of interviews with school teachers. We analyzed the results through framework analysis and in the light of scientific literature. This paper reports on results of this study. We identified a list of requirements for learning tools that bridge this gap between in-class and field learning.

## 2 Literature Review

10 years ago, a review of field learning strategies was presented in [5]. The authors identified the following key factors for effective field trips: 1) longer outdoor

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 373–378, 2014.
© Springer International Publishing Switzerland 2014

experiences, 2) well designed preparation and follow-up work, 3) use of a range of learning strategies, 4) emphasis on the role of facilitation for the learning process, and 5) a close link between field trip aim and practice.

With respect to deploying mobile learning technology on field trips, the majority of tools that have been reported on in literature focused on facilitation in a specific learning domain (e.g., history [3] or social studies [4]) In addition, researchers focused on designing technology-based facilitation for incorporating a particular learning theory or teaching method (for instance, serious games) into mobile learning. Wu and fellows [3] presented a system intended to help fifth grade students with learning historical and cultural contents through treasure hunting activities during a field trip. Chee and fellows discussed a similar approach to help high school students (approx. aged 15 years) with citizenship learning through a game-based activity [4]. Both examples demonstrated the effectiveness of serious games in enhancing learners' engagement in the given situations and domains (social studies), but these tools might not be helpful in other domains and for other field learning scenarios.

Vavoula and fellows [1] took a more flexible approach for facilitating field trips. They developed a system that allowed learners to gather data through mobile devices while information was automatically sent to a website which the learners could review later. The system facilitated various stages of field trips and could be used in various educational domains. Although they provided the flexibility for incorporating other computers through a website, the potential of in-class technologies was not fully exploited in the approach. Giemza and fellows [2] presented a system for supporting field trips. The system facilitated various stages of field trips (preparation, on-site data collection, and reflection after the field trip) and incorporated an authoring interface that allows for defining certain trips with specific tasks, student groups, devices and locations. This system followed a generic approach for the technical support for field trips with mobile devices and personal computers, but the role of in-class technologies (e.g., electronic whiteboards) was not explicitly considered in the approach.

In summary, many existing learning environments aiming at facilitating field trips were either focused on supporting certain on-site activities of field trips, or were specialized for a particular domain. Moreover, in-class technologies were not fully integrated with mobile devices which prevent a seamless flow of learning activities before and after field trips. This was the motivation for us to reexamine field trip activities and to identify their technology-related requirements.

## 3 Qualitative Study

The goal of this study was to investigate the key parameters of field trips, teaching strategies, learning goals, and the role of technology from the teachers' perspective.

### 3.1 Design

The study was comprised of semi-structured face-to-face interviews with 10 teachers: a basic protocol with general questions was given, but in some cases, more specific questions were required for further discussion. All the participating teachers employed field trips regularly. The teachers came from four schools in two states of Germany. Eight of them were female and two were male. Their teaching experience ranged between 6 and 32 years. The domains taught by these teachers ranged from natural sciences to social sciences. The profiles of teachers are summarized in Table 1. All teachers volunteered to take part in the interviews, which took between 20 to 30 minutes. The interviews were audio recorded and later transcribed. Interviews were conducted in the teachers' respective schools by the first author of this paper. Data was analyzed through framework analysis. Data was coded using textual codes and categorized into thematic framework (charting).

**Table 1.** Profile of teachers participating in the study

| No. | Gender | Experience | Teaching Disciplines | Grades |
|---|---|---|---|---|
| 1 | Female | 13 years | English, Geography | 9 to 12 |
| 2 | Female | 13 years | Physics, Mathematics | 9 to 12 |
| 3 | Male | 14 years | Computer Science, Sports | 9 to 12 |
| 4 | Female | 10 years | Computer Science, Mathematics | 6 to 12 |
| 5 | Female | 20 years | German, English, French | 5 and 6 |
| 6 | Female | 16 years | German, French | 5 to 11 |
| 7 | Female | 32 years | Biology | 5 to 12 |
| 8 | Male | 6.5 years | German, Political Studies | 5 to 12 |
| 9 | Female | 19 years | Chemistry, Biology | 5 to 12 |
| 10 | Female | 18 years | Chemistry, Biology | 5 to 12 |

### 3.2 Key Issues Identified by Framework Analysis

The main themes that emerged from the analysis of data are as follows:

1. **Nature of the field trips.** The nature of the trips depended strongly on the relevant subject, grade and teacher's disposition and enthusiasm and on available resources. On average, the teachers who participated in this study had two or three field trips per academic year in their respective fields. In case of visits to local places, the excursions normally took between a half and a full day. Intra-city field trips took from one to three days. Abroad excursions were generally longer in duration, ranging from seven to fourteen days. The main theme or learning objectives of a particular field trip were always related to the school curricula.
2. **Visited places.** The common places for school excursions were museums, exhibitions, observatories, universities' laboratories, data centers, and work places (e.g., factories and mines), places of geographical, ecological and historical significance such as botanical gardens or historical land marks. Local and intra-city trips were more common and frequent than excursions abroad.
3. **Pre-trip preparation.** All field trips were succeeded by pre-trip preparations. The preparation included introductory lectures, distribution of helping materials, trip agendas and work sheets, and discussions. The extent and duration of these pre-trip

activities varied depending on the nature of intended trip. These activities ranged from one lecture session to several weeks of preparations.

4. **On-site activities.** In general, the activities during a field trip could be divided into two categories: *Observation* and *Exploration*. In museums, aquaria, observatories, and production facilities, learners followed the observation strategy. *Observation* was generally facilitated by detailed descriptions given by a teacher or a site representative. Learners collected data in form of pictures and videos using their mobile devices. On the other hand, in the *exploration* mode, learners actually performed some experiments on their site of visit, e.g., laboratories and natural habitats of particular plants or animals. These activities included learning to use the equipment and facilities available in laboratories, collecting samples of particular species or objects, and taking various atmospheric measurements.
5. **Teaching strategies.** Teachers employed several teaching strategies to engage learners in outdoor activities: *teacher-guided* and *unguided* visits. During a *teacher-guided* visit, teachers provided direction as well as well assistance during the whole course of excursion. During excursion, in addition to provided descriptions, teachers encouraged learners to read labels, to refer to helping materials provided by teachers or information (paper or multi-media) provided by the site representatives. On the other hand, *unguided* visits were more flexible in nature. Learners followed observation or exploration strategies but received little or no guidance from teachers. For instance, teachers merely pointed out the resources, but the learners were encouraged to complete a task by researching themselves. Teacher-guided trips were conducted more frequently than unguided trips.
6. **Follow-up activities.** In most cases, follow-up activities were held after the field trips. Learners had homework assignments, in-class discussions or presentations of their experiences during the field trips. Some teachers emphasized more on post-trip activities than others. In some extreme cases, teachers arranged field trips either very early or very late in curricula. In case of an early trip, teachers referred back to a particular field trip as they advanced in the course (e.g., Teacher No. 1), while a late trip generally had extensive pre-trip activities but limited or none follow-up work (Teacher No. 3).
7. **Evaluation.** The most commonly employed technique of evaluation after a field trip was providing a homework assignment or an in-class group task. Peer review had also been employed by some teachers. In some cases, no particular evaluation was carried out after a field trip (Teacher No. 2, 4 and 5) - even though, as stated, the trip did have a curricular relevance and were not just aside activities.
8. **Learning goals.** All participants of this study agreed that field trips had multi-fold impact on the whole learning process. One main objective had been enhancing learner's understanding of a particular theme by providing learners "direct" and "first-hand" knowledge (expressed by teachers 1, 7, 8, and 9) and helped them to create a link between knowledge gained in class and in real world (Teacher No. 6). Teachers had also observed that field trips had a positive effect on learners' behaviors and on their social skills and communication abilities (Teacher No. 3 and 6). Most of the teachers also made sure that the learners had fun during their trips.

9. **Collaborative learning.** A very important aspect of a field trip is collaborative learning. All teachers reported that learners worked in groups not only for on-site activities, but also for preparation and follow-up activities. They engaged in group discussions, reviewed, re-arranged their data together, and presented their findings in groups. These pre and post-trip activities always used some form of technology (e.g., desktop/laptop, electronic whiteboards). Teachers reported that for in-class activities, learners generally worked in groups of 3 or 4 participants, but for on-site activities the group size varied between 3 to 8.
10. **Computer support for field trips**. Teachers and learners both utilized various technologies in various stages of field trips. They searched resources on the web, collect data using their mobile devices, used computers to review and re-arrange their data, and presented their work on a projector. But there was usually no good organization or scaffolding of all these activities. Another factor was the teachers' preference. While some teachers encouraged the use of technology, others pointed out some disadvantages associated with the use of devices. Major concerns included learners' distraction from the intended learning goal, students' lacks of skills regarding a particular system and concerns about negative impacts of computer use on some skills, e.g., grammar in case of language (Teacher No. 1), and about additional workload in form of maintenance of devices (Teacher No. 5).

## 4 List of Requirements for Learning Tools

The themes that emerged from the interview data show that the practice of field trips largely followed established strategies, but the way of employing learning technologies was not consistent. Teachers and learners took advantages of learning technologies but that practice was largely based on individual preference, and there was little organization or integration of the employed tools. While technology is available to facilitate both formal and informal learning, we have drawn a list of requirements for learning tools intended to support field learning.

1. **Specialized support for each stage of field learning.** As field trips are integrated into formal learning and take places in various locations, specialized support is required depending on physical locations, nature of the learning tasks, underlying learning theories, number of participants, and intended activities. A single technology might not be able to address all these issues in any given situation. Rather, an integrated approach drawn on various technologies would certainly be beneficial where the particular task is facilitated by most suitable technology - i.e., devices with larger displays for group work, mobile devices for on-site data collection, and personal computers (laptops and tablets) for individual assignments where keyboard input is essential e.g., essay or report writing.
2. **Seamless integration.** The incorporation of multiple technologies needs to focus on providing facilitation without introducing additional complexity. That demands for an easy incorporation of new devices, automatic synchronization and easy data sharing between all devices along with the added value of devices' features.
3. **Support for multiple learning theories and strategies.** The literature review and our study results showed that field trips often comprised multiple combined learning strategies, so learning tools should be flexible in this regard.

4. **Support for administrative activities.** Distribution of helping materials, worksheets, plans for on-site activities, and trip schedules are essential tasks for effective field trips. Such administrative tasks might have little pedagogical underpinning but still play a vital role in both formal and informal learning and should be considered in facilitation systems for field learning.
5. **Teachers' control and preference.** The tools should also allow teacher's overruling and give him flexibility. As the teacher is the main facilitator in traditional school field trips, a teacher's judgment on when and where and to which extent learning tools should be used is very important.
6. **Evaluation support.** Learners' assessment is an important aspect in formal learning. Field trips comprise features of both formal and informal learning, and a specialized support for learning assessment should be provided by learning tools used in field trips. This is not possible with existing tools.
7. **Ease of use and training.** The learning tools should be easy to learn and easy to use. Ease of use is a general requirement for any interactive system but it is especially critical for learning tools so that the student's focus is on the themes of the trip and their learning, and not on finding out how to use the tools.

## 5 Conclusion

This paper presented the outcomes of a qualitative study that aimed at finding requirements for learning tools to facilitate learning both in conventional classroom settings and on field trips. We conclude that facilitation tools for field learning should incorporate heterogeneous technologies, provide seamless integration between various technologies, be flexible to incorporate various learning methodologies and preferences, provide control to teachers, support evaluation, and be easy to learn.

## References

1. Vavoula, G., Sharples, M., Rudman, P., Meek, J., Lonsdale, P.: Myartspace: Design and evaluation of support for learning with multimedia phones between classrooms and museums. Computer & Educaton 53, 286–299 (2009)
2. Giemza, A., Bollen, L., Seydel, P., Overhagen, A., Hoppe, H.U.: LEMONADE: A Flexible Authoring Tool for Integrated Mobile Learning Scenarios. In: 6th IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education, pp. 73–80. IEEE Computer Society (2010)
3. Wu, S., Chang, A., Chang, M., Yen, Y.R., Heh, J.S.: Learning Historical and Cultural Contents via Mobile Treasure Hunting in Five-harbor District of Tainan, Taiwan. In: 6th IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education, pp. 213–215. IEEE Computer Society (2010)
4. Chee, Y.S., Tan, E.M., Liu, Q.: Statecraft X: Enacting Citizenship Education using a Mobile learning Game Played on Apple iPhones. In: 6th IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education, pp. 222–224 (2010)
5. Rickinson, M., Dillon, J., Teamey, K., Morris, M., Choi, M.Y., Sanders, D., Martin, P.B.: A review of research on outdoor learning. National Foundation for Educational Research and King's College (2004)

# Are We All on the Same Boat? Coordinating Stakeholders for the Design of MOOCs

Carlos Alario-Hoyos, Mar Pérez-Sanagustín, and Carlos Delgado Kloos

Department of Telematic Engineering, Universidad Carlos III de Madrid
Av. Universidad, 30, 28911, Leganés, Madrid, Spain
{calario,mmpsanag,cdk}@it.uc3m.es

**Abstract.** One of the gaps that arises from the recent emergence of Massive Open Online Courses (MOOCs) is the lack of methodologies, tools and models for supporting the instructional design of these complex courses, which typically involve several stakeholders (e.g., teachers, audio visual technicians, institutional staff…). One of the few approaches addressing this gap so far is the MOOC Canvas, a framework for supporting the description and design of MOOCs. This paper presents a first approach towards a methodology that applies the MOOC Canvas as an instrument for coordinating the needs and interests of the different stakeholders involved in the preparation and enactment of MOOCs. Also, this paper provides insights about the use of this methodology with different stakeholders in three workshops. The results of the workshops show a positive perception of the methodology and of the use of the MOOC Canvas as the main instrumental support.

**Keywords:** MOOCs, instructional design, MOOC Canvas, stakeholders.

## 1 Introduction

MOOCs (Massive Open Online Courses) have attracted much attention from the educational community thanks to initiatives such as edX, Coursera, FutureLearn or MiríadaX. Now, many teachers and institutions are facing the challenge of running MOOCs not to fall behind. Nevertheless, preparing and enacting a MOOC is much more complex than preparing and enacting a traditional (online) course, mainly because of two reasons: the need for adopting new technologies (including video-based technologies), and the massive number of participants that these courses can reach. MOOCs usually involve several stakeholders besides teachers, such as audio-visual technicians, institutional staff, library staff and system administrators, among others, who collaborate in the preparation and enactment of the MOOC. The large number of stakeholders in MOOCs requires extra coordination while the course is running, and especially, while designing the course.

Despite the particularities of MOOCs, and the large number of these courses available nowadays on the Web, only few works address their instructional design [1, 2, 3, 4, 5, 6]. Most of these works offer recommendations and/or best practices for designing and setting up MOOCs from the teachers' perspective. However, these works do

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 379–384, 2014.
© Springer International Publishing Switzerland 2014

not consider in their proposals the need for coordination among the stakeholders that requires the instructional design of MOOCs.

This paper presents a methodology to guide and coordinate all the stakeholders in the instructional design of MOOCs. This methodology builds upon the MOOC Canvas [7], a framework for supporting the description and design of MOOCs inspired by the Business Model Canvas [8]. The MOOC Canvas is a visual, high-level representation of the MOOC that helps stakeholders discuss and reflect on eleven interrelated issues through a set of driving questions. The methodology has been tested with representatives of the main stakeholders involved in a MOOC in three workshops.

The remainder of this paper proceeds with a brief description of the MOOC Canvas in Section 2. Section 3 presents the methodology that builds upon the MOOC Canvas. Next, section 4 shows the main insights about the instructional design of MOOCs extracted from the three workshops, and Section 5 draws the conclusions and future lines derived from this work.

## 2 The MOOC Canvas

The MOOC Canvas[1] is a conceptual framework to support the description and design of MOOCs [7]. This framework offers a visual and understandable guidance during the instructional design stage of a MOOC, facilitating the coordination and discussion of the eleven most important issues conditioning its design (see Figure 1). These issues are organized into **available resources** (issues 1-4) and **design decisions** (issues 5-11). The MOOC Canvas should be completed following the issue numbers and reflecting on the driving questions associated to each issue.

The **available resources** refer to the key resources available for setting up and running the MOOC. These resources include: (1) *Human* resources, which are the people that can commit to take part in the MOOC (considering all the potential stakeholders); (2) *Intellectual* resources, which are the existing learning contents and other related materials that can be reused (or adapted) to build the MOOC; (3) *Equipment* resources, which are the hardware and software resources available for generating the MOOC contents; and the (4) *Platform* in which the MOOC will be deployed, and the features this platform offers. Since teachers may be aware of some, but not all of the available resources, this part of the MOOC Canvas should be completed in coordination with other stakeholders. Particularly, teachers need to coordinate with institutional staff to agree on the available human resources; with library staff to determine the available intellectual resources; with audio visual technicians to know the available equipment resources; and with system administrators to know the (remote or locally hosted) platform in which the MOOC will be deployed, and the features this platform offers. This part of the MOOC Canvas must be completed before starting making design decisions about the MOOC, since such decisions will be strongly constrained by the available resources [7].

The **design decisions** refer to the main decisions teachers should make when designing a MOOC, taking into consideration the constraints imposed by the available resources, as previously established. These design decisions include: (5) the *General*

[1]MOOC Canvas Website: `http://www.mooccanvas.com`

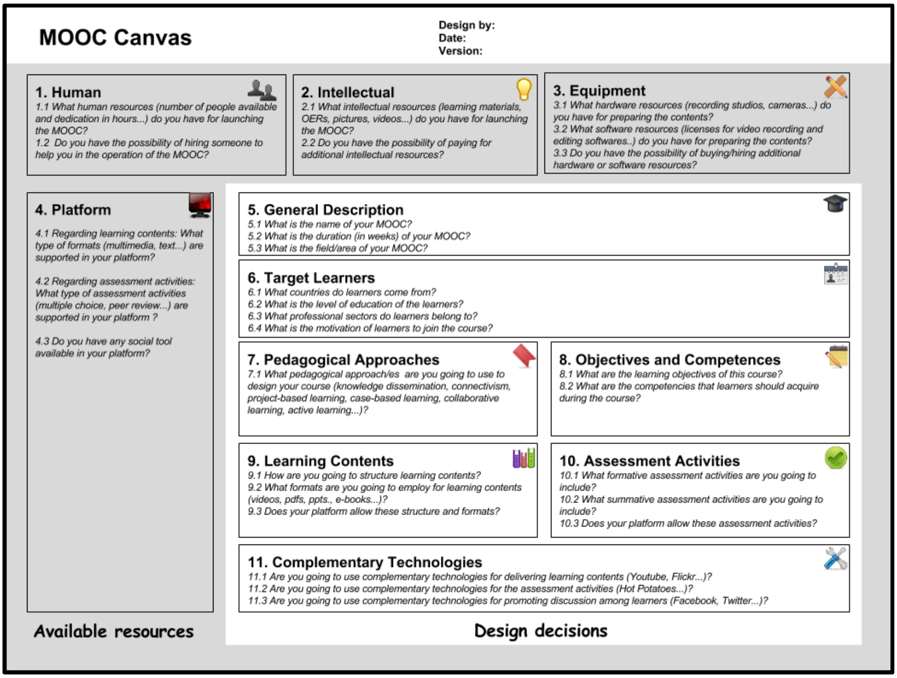



**Fig. 1.** The MOOC Canvas

*Description* of the MOOC (name, duration, area); (6) the *Target Learners* that are expected as the potential audience; (7) the *Pedagogical Approaches*, didactics or concrete teaching methods to be used during the MOOC enactment; (8) the *Objectives and Competences* that are expected to be acquired by the participants; (9) the *Learning Contents* that will be provided and their delivery formats; (10) the *Assessment Activities* that will be included in the MOOC (either formative and summative); and (11) the *Complementary Technologies* that are not directly provided by the selected MOOC platform, but that are still needed to meet the remaining design decisions. A MOOC can be delivered by several teachers and so, design decisions should be discussed and agreed among them, sometimes even with the support of other education experts in instructional design. The MOOC Canvas represents a useful tool for capturing and understanding at a glance the overall design decisions taken.

## 3 A methodology for the instructional design of MOOCs

This methodology builds upon the MOOC Canvas (as the main instrumental support) and includes three phases (A, B and C), each of which involves several steps (see Figure 2). Phase A is for setting the available resources, Phase B is for making the design decisions, and Phase C iterates on specific steps of Phases A and B. The same version of a MOOC Canvas is the instrument shared among all the stakeholders along the three phases for coordination purposes.

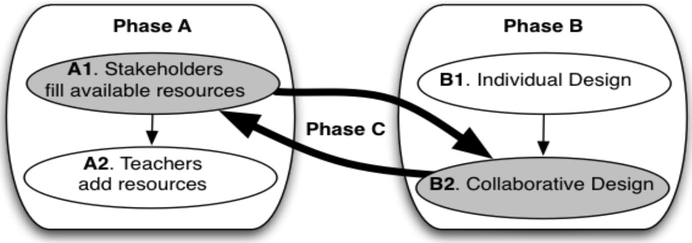



**Fig. 2.** Methodology for the instructional design of MOOCs

The first step in Phase A (A1) is that the different stakeholders partially fill in the available resources, according to the information they have. Typically, but depending on each organization, institutional staff fill in issue 1 (*Human*); library staff fill in issue 2 (*Intellectual*); audio visual technicians fill in issue 3 (*Equipment*); systems administrators fill in issue 4 (*Platform*). In the second step (A2), the teachers individually add to the MOOC Canvas the resources they can provide (e.g., time commitment, existing materials, etc.).

In the first step of Phase B (B1), teachers individually design the course they want to deliver taking into account the available resources. In a second step (B2) all the teachers participating in the MOOC meet together (face-to-face or online) for discussing and agreeing on the design decisions (issues 5-11), based on their individual designs and the available resources. An education expert in instructional design may support this step (if possible). A first full version of the MOOC Canvas is obtained at the end of this second step.

In Phase C, all the stakeholders will iterate over the second step of Phase B (B2) and the first step of Phase A (A1) to produce new versions of the MOOC Canvas. These new versions will capture the need for new resources that may be detected during teachers' group discussions (i.e., more manpower, new equipment, etc.) (B2). These needs will be communicated to the corresponding stakeholders, who will check the MOOC Canvas to find out if these needs are properly justified. If so, and if the resources can be acquired, the stakeholders will add them to the MOOC Canvas (A1). Otherwise, teachers should meet together again to find alternatives to their design decisions (B2). This third phase is repeated iteratively until reaching an agreement on a final MOOC Canvas. The final MOOC Canvas can be kept by the institution to assess the quality of the instructional design of the MOOC and verify its fulfilment.

## 4 Application of the Methodology with Different Stakeholders

This methodology has been tested with different stakeholders (n=27) in three workshops. The first workshop was conducted in June 2013 with the participation of 9 educators, experts on instructional design, from the Universidad de Cádiz (Spain). The second workshop was conducted in December 2013 and involved 8 participants from the Universidad Carlos III de Madrid (Spain) with different roles: library staff, audio-visual technicians and systems administrators. The third workshop was conducted in January 2014 and involved 10 participants, experts on instructional design, from the SRI International, a nonprofit, independent research and innovation center based on Northern California (USA). During these three workshops, participants worked in small group (3-5 people) on a MOOC Canvas using pen and paper to design a MOOC of their choice following Phases A and B. Participants could also follow Phase C to iterate on steps A1 and B2 in order to refine the MOOC Canvas. The *Platform* (and its characteristics) was set in advance by the workshop organizers as the resource available, but the participants completed the remaining issues. See more details about the WS in [7].

The outcomes of these workshops were a voluntary anonymous questionnaire filled out by 20 of the 27 participants, and the MOOC Canvas generated collaboratively in small groups. The questionnaires revealed a complete agreement among all the participants that *the MOOC Canvas is a good discussion tool in the instructional design of MOOCs*. This is a significant result since the samples of people in these three workshops involved different stakeholders expected to participate in different phases of the methodology. The experts on instructional design gave positive feedback about the order and structure in the design of the MOOC, as well as on the resulting MOOC Canvas for visually representing the overall instructional design. The remaining stakeholders gave positive comments to the organization of the design, and its value for thinking and reflecting before making decisions. The time required to fill in the MOOC Canvas (and by extension to follow the methodology) was the main negative comment. Also, participants highlighted that this methodology requires start working together on the MOOC design as soon as possible, which is not always possible.

Figure 3 presents an example MOOC Canvas generated by both audio-visual technicians and systems administrators in the second workshop. The workshop organizers established that this MOOC should be deployed in edX, which is a platform these

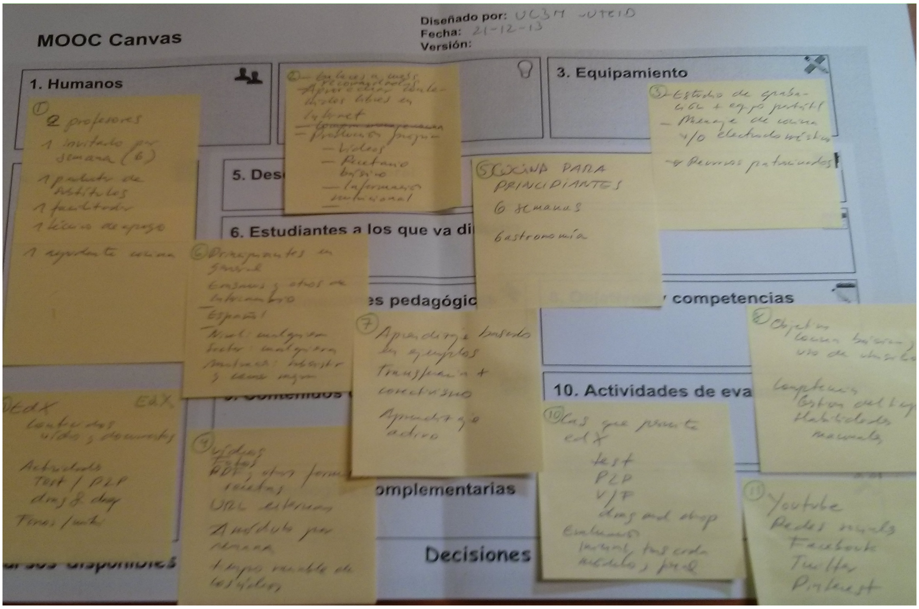



**Fig. 3.** Example of MOOC Canvas filled in in one the workshops

technical staff were familiar with, but there were no additional restrictions about the remaining available resources. With this activity, these stakeholders, who are not used to take into account the pedagogical issues underlying MOOCs, could step into the shoes of teachers making design decisions. This is a way to facilitate the coordination and understanding between teachers and these other stakeholders.

## 5 Conclusions and Future Work

This paper has presented a methodology built upon the MOOC Canvas to facilitate the coordination among the different stakeholders that participate in the preparation and enactment of MOOCs during their design. This methodology has been tested in three workshops with 27 stakeholders, who performed the complete cycle of the methodology. Results show a positive perception of the methodology and of the use of the MOOC Canvas as the main instrumental support.

Future lines of work include: a) using the methodology in other courses and with other stakeholders to refine it and improve the MOOC Canvas, and b) promoting the institutional adoption of this methodology in our University, combining it with other MOOC design best practices and recommendations [6]. Finally, since the MOOC Canvas is currently a Google Drawing document, we are already working on developing a software application for sharing the different versions of the MOOC Canvas and facilitating the remote coordination among the different stakeholders when designing MOOCs.

**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. Grover, S., Franz, P., Schneider, E., Pea, R.: The MOOC as Distributed Intelligence: Dimensions of a Framework & Evaluation of MOOCs. Technical Report, Stanford (2013)
2. Jermann, P., Bocquet, G., Raimond, G., Dillenbourg, P.: The EPFL MOOC Factory. In: Proceedings of the European MOOC Stakeholder Summit 2014, pp. 228–233 (2014)
3. de Waard, I.: The MOOC Guide, http://moocguide.wikispaces.com
4. de Waard, I.: MOOC Yourself - Set up your own MOOC for Business, Non-Profits, and Informal Communities (2013)
5. Building a Coursera Course. Technical Report, Duke Center for Instructional Technology (2013)
6. Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Gutiérrez-Rojas, I., Leony, D., Parada, G.H.A.: Designing your first MOOC from scratch: recommendations after teaching 'Digital Education of the Future'. eLearning Papers, From the Field 37(3), 1–7 (2014)
7. Alario-Hoyos, C., Pérez-Sanagustín, M., Cormier, D., Delgado-Kloos, C.: Proposal for a conceptual framework for educators to describe and design MOOCs. Journal of Universal Computer Science 20(1), 623 (2014)
8. Osterwalder, A., Pigneur, Y.: Business Model Generation. John Wiley & Sons, Inc., Hoboken (2010)

# Reliability of Web-Based Information in Inquiry Projects

André Luis Corrêa[1], Cíntia Regina Lacerda Rabello[2], and Alexandra Okada[3]

[1] Universidade Estadual Paulista, Brazil
andrelc@fc.unesp.br
[2] Universidade Federal do Rio de Janeiro, Brazil
cintia_rabello@hotmail.com
[3] The Open University, UK
a.l.p.okada@open.ac.uk

**Abstract.** This case study aims at identifying how a community of secondary school students selects web-based information and factors associated with the reliability of online reference sources during their collaborative inquiry (co-inquiry) projects. This study, conducted in a public secondary school in Brazil, focused on information literacy skills for collaborative open learning (colearning). The research is based on qualitative content analysis implemented on the online platform weSPOT. Although students are mindful of the importance of comparing different sources of information they seem not to be aware of reliability in online environments. Teacher's guidance is essential to support co-learners in developing competences, particularly related to critical thinking.

**Keywords:** Secondary school information literacy, co-inquiry, weSPOT.

## 1 Introduction

Due to the rapid growth of user-generated content, information reliability is considered one of the key skills for digital competence [3]. Educators play a key role in helping learners become aware of issues around the validity and reliability of information available on the web as well as the legal and ethical principles related to the interactive use, remix and redistribution of information.

Various studies highlight that online sources have been used more frequently for research tasks by students than traditional printed sources [5][11]. Informal online sources have been expanding dramatically: Wikimedia (12 million files), Facebook (3.5 billion msgs), Twitter (1 billion msgs) and YouTube (604,800 hours of video). Academic online resources have also increased through various OER, MOOC and digital libraries [6].

The aim of this paper is to investigate how secondary school students select web-based information and identify factors associated with the reliability of information sources during collaborative inquiry (co-inquiry) in weSPOT. The weSPOT project (2013-2015) is a working environment with social, personal and open technologies for inquiry based learning funded by the European Commission. This case study is the first pilot of weSPOT in Biodiversity from October to November 2013.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 385–390, 2014.
© Springer International Publishing Switzerland 2014

## 2 Colearning and Inquiry-Based Learning

The concept of colearning emphasizes the role of teachers and students as both collaborative partners in the process of learning [8]. According to Freire [4] learners have their existent knowledge and previous experiences that constitute important contexts for educators to learn and improve opportunities for teaching by connecting their existing context with new knowledge. Colearning acts toward student centred learning as well as building a more genuine "community of practice" for collective construction of knowledge [1]. This concept has become more popular due to the rapid advances of Web 2.0, which allows the creation and exchange of user-generated content, OER and social networks. Due to the philosophy of openness, the process of colearning (collaborative open learning) is enriched through wide participation for creating, adapting and reusing content [6].

The concept of co-inquiry – cooperative inquiry – is based on the terms "community of inquiry" [2][10] and "inquiry-based learning". It refers to a process of building knowledge through collaborative investigation where colearners are engaged in developing arguments based on evidence [9]. Through inquiry students are guided to develop knowledge and scientific ideas, as well as an understanding of how scientists study the world. They are encouraged to make observations; create questions; search content in various references. Students who develop their skills as co-investigators are able to use evidence to explain questions; use tools to collect and interpret data; describe the process including outcomes and communicate their findings.

Digital natives are immersed in a world of digital technologies, but it does not necessarily mean they are digitally or information literate. When we consider these learners it is important to understand how they search for information on the web and how they evaluate the information they find as to work on helping them develop their digital and information literacy and become more effective learners and achieve better results in their colearning and co-inquiry tasks. This study focuses on "how learners analyse reliability of web-based information in their inquiry through weSPOT".

## 3 Methodology

This case study was conducted with twelve 15 to 17-year-old secondary students who are in the first year of secondary school in Bauru, São Paulo, Brazil. In order to investigate how students search and select information sources on the web the study was divided into two different phases which comprised the use of different research instruments for qualitative data analysis: (1) A structured inquiry project using the weSPOT: online forum, questions and data collection; (2) a semi-structured survey.
Students used their own mobile phones through the application ARLearn to take pictures about the school garden, which were automatically shared in weSPOT platform and analysed collaboratively in the forum discussion. They started their co-inquiry project with some questions:

- How many living species are present in the photo you took?
- Which abiotic factors influence the presence of these living species?

- Supposing you do not know the meaning of species and/or abiotic factors, how would you search it so that you could answer the previous questions?
- How would you confirm that the information you collected is reliable?

Data from weSPOT,(images and discussion) as well as the survey (reliability of web-based information in inquiry projects) were analysed based on the "C" model: Competences for Colearning and Co-inquiry [7] as it is shown in figure1.

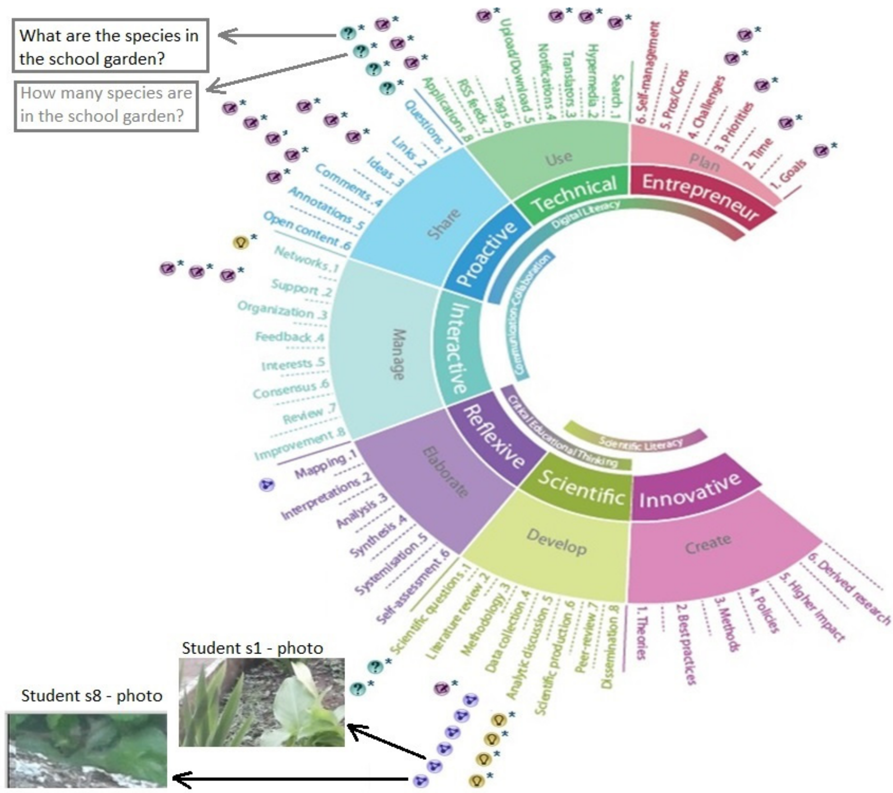



**Fig. 1.** "C" model analysis - Competences for Colearning and Co-inquiry with qualitative data from weSPOT (ARLearn images and forum discussion messages) UR: colearn.open.ac.uk/c

## 4 Results

The semi-structured survey revealed that eleven students use social media and networks for studying. Platforms such as blogs, webconferencing tools and Virtual Learning Environments (VLE) are never used by the students. In relation to the devices for studying, twelve claimed using desktop computers rather than mobile devices (mobile phones or tablets). Although eight students declared having mobile phones and, 4 of them have access to tablets; all of them claimed using only computers for studying in the school or at home. When asked to use the weSPOT platform along with the ARLearn app, students showed no difficulty in carrying out the task.

In terms of the key competences for colearning and co-inquiry [7], ten students claimed establishing objectives to be achieved when using the web for their studies and identifying benefits and difficulties to think about other alternatives. However, only two students affirmed having initiative to establish the next steps towards self-learning. Twelve students have claimed making use of many different web tools, such as search engines, hypermedia, translators, and notifications, downloads/uploads and software. Tags and RSS feeds were, however, pointed as of no use by the students. These two tools are particularly important in terms of information literacy, as they are useful to retrieve and filter information on the web, facilitating and enhancing students' learning. Eleven students claimed sharing questions and doubts, web links (images, websites, and videos), comments, annotations and open content in online environments. These actions show that learners are used to interacting in the web. Seven students declared rarely sharing ideas and suggestions; which also corroborates the idea of an educational system still based on content transmission rather than sharing of experiences and ideas as well as critical and creative thinking. Seven students also claimed using the web to search for information they do not know through web tools such as Google, Bing, Yahoo and YouTube. Moreover, eleven students claimed comparing content from different webpages. However, it was noticed that students have difficulty in explaining how they evaluate the reliability of the information they used in their inquiry, showing little information literacy skills.

Students were also asked which websites they often use and why. Five students claimed using Wikipedia as their most common information source, while the others mentioned using Google and Yahoo search engines and YouTube. As for the reasons for using these sources, students mentioned popularity and clear content as important factors for their preference of information sources. It was possible to notice students' perception of Wikipedia as a reliable source of information for their studies. Despite being a controversial issue, Wikipedia may be a reliable source when supported by other search sources.

Nine students claimed using more than three different sources for comparison purposes. From the answers above we can infer that students use quite a few information sources in their searches related to the verification of the information they will use in their co-inquiry projects. They are aware of the importance of comparing information sources in order to verify their credibility and reliability. However, when asked about how they evaluate that the information they find is appropriate for their study and research projects, students were unable to establish reliability criteria mentioning comparison of different sources as the main reliability factor they use as it can be seen in the following extracts: *"I check if the same information is available in different websites and confirm the sources"(S1); "I read at least two websites and check if the information matches"(S5); "I try to understand if the topic is in the context that was requested. If it is I continue, if it's not, I search another website"(S3).*

Just one student claimed following the teacher's guidelines and requests. Although most students use web materials for their search, we also asked them if they use any other information sources. Most students claimed using course books or books they have at home. They also use books that are recommended by their teachers or a family member. None of them use the school library for searching for books and one student mentioned not using books at all. Besides books, students showed little use of other materials or resources such as magazines and newspapers with just a few students mentioning using these resources. All the students compared different information sources such as websites, teachers' notes and books, and some repositories. Most of them claimed comparing sources in order to find similarities. Others affirmed conducting personal evaluation of different sources or relating the question they have to the answer they find. However, most answers seemed vague and rather imprecise as it can be noticed in the following extracts: "*I compare the answers and if they are coherent to the questions and if they have the same information in more than one website*" (S3); *"I try to check if the information coincides."* (S6).

Finally, we asked how these students evaluate information quality so that they would use them or not in their projects. Eleven students indicated making personal evaluation of the information sources based on criteria such as completeness, understandability and adequacy to their needs. Students also mentioned using reliable sources and confirming the information with a teacher or a knowledgeable person.

Based on the results above, we can conclude that despite being aware of the importance of comparing different information sources on the web, secondary students lack information literacy skills in terms of recognizing and evaluating the quality of these sources. Although they do mention making use of reliable sources they were unable to explain the criteria/factors that make an information sources reliable. Despite comparing more than three sources, they usually select the first results on a search engine page ignoring reliability factors. They also seem to use limited resources in terms of search engines and websites, preferring simpler and more straightforward contents than looking for more specialised materials and sources. They seem to trust the content they find on the web, rarely recurring to books or experts to compare and check the veracity of the information they find.

The survey and weSPOT platform show that co-inquiry might be useful for:

- supporting students to manage next steps towards social and self-learning by personalised inquiry phases and learning analytics about their participation;
- encouraging colearners to share more ideas and suggestions through weSPOT widgets: pinboards, wikipages, mindmaps, mobile data collection;
- adding votes and rates to colleagues' contributions and peer-review comments, including assessing reliability of information sources;
- retrieving and filtering information through tags and RSS feeds.

Additionally, it was observed that teachers play an important role for:

- sharing guidelines for students to check reliability of information sources;
- encouraging students to check and assess references;
- using advanced search engines and collaborate with an open repository;
- helping students combine various sources of reliable information.

Based on the new functionalities, which will be available soon in weSPOT, during the next pilot study, educators and researchers aim to encourage colearners to develop their scientific explanations with peer-reviews and create new inquiry projects with high levels of autonomy for a collaborative guided inquiry.

## 5 Conclusions

This study shows that secondary students in a public school in Brazil use the web as their main tool to search for information. Even though students claim comparing contents as to evaluate the quality of web-based information, this attitude seems ineffective as they present few abilities in evaluating information sources, which seems to be connected to preconceived knowledge, i.e., some information previously obtained that might not be reliable, hindering students' information acquisition and learning.

Although these students are frequent social network users and natural colearners, they do not consider their social networks for studying/ learning activities. The study on skills for 21st century learning may be of great importance as it is not conditioned to the acquisition of school content, but to the development of new competences such as critical-creative thinking and scientific literacy through collaborative inquiry-based learning. Educators play an essential role in supporting students to build better information search strategies. Additionally, working environments with social personal open technologies such as weSPOT might be useful for educators and colearners develop investigations on individual and collective issues, conduct efficient searches on the web, and construct knowledge collaboratively.

**Acknowledgements.** weSPOT is funded by EC FP7-2013 N° 318499 and supported by CAPES Brazil.

## References

1. Brantmeier, E.J.: Empowerment Pedagogy: Colearning and Teaching. Indiana University (2005)
2. Dewey, J.: Logic: The Theory of Inquiry. Hold Rinehart and Winston, New York (1938)
3. European Communities. The Key Competences for Lifelong Learning – A European Framework (2007)
4. Freire, P.: Educaçãocomoprática para a liberdade, 17$^a$ edn. Paz e Terra, Rio de Janeiro (1986)
5. Lee, J., Paik, W., Joo, S.: Information resource selection of undergraduate students in academic search tasks Information Research, 17(1) paper 511. National Academy Press (2012)
6. Okada, A., Mikroyannidis, A., Meister, I., Little, S.: "Colearning" – Collaborative Open Learning through OER and Social". In: Okada, A. (ed.) Open Educational Resources and Social Networks. EDUEMA, Sao Luis (2012)
7. Okada, A.: Competencias-chave para coaprender: fundamentos, metodologias e aplicações. DeFato (2014)
8. Smith, F.: Joining the literacy club. Heinemann, Portsmouth (1988)
9. Smithenry, D.W., Gallagher-Bolos, J.A.: Whole-Class Inquiry. NSTA, Virginia (2009)
10. Shields, P.: The Community of inquiry: Classical Pragmatism and Public Administration. Administration & Society 35(5), 510–538 (2003)
11. Xie, I., Joo, S.: Selection of information sources: types of tasks, accessibility and familiarity of sources. Proceedings of the American Society for Information Science and Technology 46(1), 1–18 (2009)

# Making Sense of Bits and Pieces: A Sensemaking Tool for Informal Workplace Learning

Sebastian Dennerlein[1], Matthias Rella[4], Vladimir Tomberg[3], Dieter Theiler[4], Tamsin Treasure-Jones[2], Micky Kerr[2], Tobias Ley[3], Mohammad Al-Smadi[3], and Christoph Trattner[4]

[1] Knowledge Technologies Institute, Graz University of Technology, Austria
{sdennerlein,matthias.rella}@tugraz.at
[2] University of Leeds, UK
{t.treasure-jones,m.p.kerr}@leeds.ac.uk
[3] Tallinn University, Estonia
{vtomberg,tley,msmadi}@tlu.ee
[4] Know-Center, Austria
{dtheiler,ctrattner}@know-center.at

**Abstract.** Sensemaking at the workplace and in educational contexts has been extensively studied for decades. Interestingly, making sense out of the own wealth of learning experiences at the workplace has been widely ignored. To tackle this issue, we have implemented a novel sensemaking interface for healthcare professionals to support learning at the workplace. The proposed prototype supports remembering of informal experiences from episodic memory followed by sensemaking in semantic memory. Results from an initial study conducted as part of an iterative co-design process reveal the prototype is being perceived as useful and supportive for informal sensemaking by study participants from the healthcare domain. Furthermore, we find first evidence that re-evaluation of collected information is a potentially necessary process that needs further exploration to fully understand and support sensemaking of informal learning experiences.

## 1 Introduction

The value of informal learning at the workplace is increasingly recognized. For example, in the healthcare domain there is a particular interest to create value out of informal experiences. However, there are numerous challenges [10]: due to high time pressure, healthcare professionals often lack time to instantly make sense out of informal learning experiences they are facing at work. Without the right support, just a few will be remembered, meaningfully connected to each other and acted upon. A huge potential for individual self-development that is now required by the English healthcare system in the yearly appraisal and five yearly revalidation process is consequently missed.

There is a wealth of informal learning experiences ranging from picking up a *toRead* whilst a short chat with a colleague to profound ones connected to a certain professional activity. Both these and many more are legitimate and account for informal learning. Respective episodic experiences are distributed over longer periods of working time and stored in episodic memory from where they need to be remembered and retrieved later

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 391–397, 2014.
© Springer International Publishing Switzerland 2014

when time permits to make sense out of them. Narrowing down and organizing corresponding experiences then involves a process of mental categorization and connection to other experiences which transforms them over time into the semantic memory. This turns the experience into more generalized knowledge that informs future action [3].

We study these activities as a sensemaking process based on the constructivist understanding of learning. Sensemaking has been subject of HCI research for decades [9][1] although mainly being studied in information gathering and intelligence tasks. It has been defined as 'the process of encoding retrieved information to answer task-specific questions' [9] consisting out of two overlapping loops [7]: information foraging for browsing and narrowing down of data & information organizing for introducing new concepts and producing final representations. Striving for a certain goal, these loops are driven by a reciprocal interchange of the generation and instantiation of respective representations to gradually fit them to the underlying data. However, this perspective excludes the individual creation of 'data' during continuous exploration as outlined by the sensemaking metaphor [2]: an individual moves through time and space, taking one step after the other, and needs to bridge gaps it encounters. Taking this view into account, sensemaking for informal learning supports the move from chaotic unstructured raw data that is created as traces for experiences to meaningful learnings that can be further shared to guide others in their actions and practices. The corresponding prototypical implementation, called 'Bits and Pieces' (B&P), is evaluated in a field setting to answer the question whether it is able to support sensemaking for informal learning.

## 2 The Bits and Pieces Prototype

Existing solutions for informal learning in the healthcare domain focus particularly on the collection of pieces of information such as Osmosis (`http://osmosis.me`) or Evernote (`http://evernote.com`). The B&P prototype we have developed and validated with healthcare professionals in several co-design sessions[1] (e.g. via participatory observation, interviews, paper prototyping) [10] is particularly designed for sensemaking of one's own learning experiences. TalkApp [8], a prototype for documenting informal experiences, which adds reflections and sustains outcomes, is related to our approach, but misses the support for organizing information bits as it focuses on the aggregation part. In contrast to this, Apolo [4] implements visual sensemaking of citation networks supported by recommendations, but lacks support for foraging of own unstructured information pieces. The B&P prototype addresses these shortcomings by leveraging a visual and adaptable approach in its design decisions which implements information foraging of one's own experiences and organizing as explicit user interface components.

B&P reflects information foraging and sensemaking by two vertically arranged widgets (see fig. 1): The upper widget (timeline view) visualizes collected notes, photos, Web links, etc. for browsing and remembering corresponding episodic learning experiences via a time-based view of bits, which can be panned and zoomed. The lower widget (organizing view) on the other hand serves for semantically organizing these

[1] The design emerged as part of the FP7 IP Learning Layers (`http://learning-layers.eu/`)

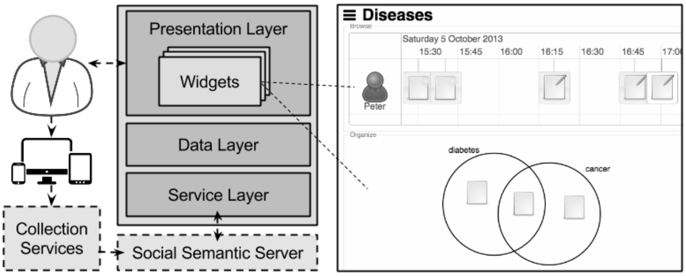



**Fig. 1.** The B&P prototype: The upper part contains the timeline widget, the lower one the organize widget. The framework is build upon RESTful Web services of the so-called Social Semantic Server loading learning bits from collection services such as Evernote.

bits by drawing circles and arranging bits like in a Venn diagram. As sensemaking is grounded in action [2], one design decision was to enable users to *interact with each of those bits*: i.e., they can be dragged and dropped from the upper to the lower widget and each bit can be selected and its detailed information viewed in a meta-data box. On account of the episodic and heterogeneous nature of informal learning [3], users need to be able to separate sensemaking according to different learning needs and easily switch between them. Hence, new sensemaking sessions can be created and switched via a dropdown menu to quickly provide anchors to the corresponding sensemaking states. The label of the current session is centrally displayed in the header of the application.

Our prototype is implemented as a *generic framework which aims at supporting the development of independent and interactive widgets* by design. It responds to the user's input immediately (through intelligent client side catching techniques) without annoying loading times to facilitate the user's direct and interactive experience. The prototype builds upon so-called social semantic services [5] in order to process and represent the unstructured informal learning bits as well as the various states of the organizing process. Henceforth, individual widgets can implement meaningful visualization techniques and interaction affordances on top of the semantics of a bit. The framework separates the concerns for remote services, data and presentation in a three-layered architecture (see Fig. 1). The service layer wraps remote Web services and prepares the bits of information whereas the data layer is responsible for curation. It keeps references between bits consistent to adaptable data integrity constraints and bundles data loading and creation logic. The presentation layer consists of one or more widgets and handles user-generated events (e.g. mouse clicks). The prototype builds upon remote (RESTful) Web services of the Social Semantic Server [5] and is implemented in JavaScript on top of VIE (`http://viejs.org`). The implementation is available on GitHub[2] as OpenSource.

## 3 Evaluation

The formative evaluation of the prototype formed part of our ongoing co-design research approach [10] involving seven professionals from two healthcare practices to

[2] `https://github.com/learning-layers/BitsAndPieces`

investigate sensemaking in the context of informal learning. The current paper reports on user tests of a first functional prototype with two diabetes specialist nurses (DSN), one healthcare assistant (HCA) and one doctor (GP). The procedure consisted of two phases to realize the necessary preconditions of informal sensemaking: (1) a two-week collection phase for gathering bits of information related to a chosen realistic learning need (e.g. discover methods for reversing diabetes or links between different kinds of dementia and diabetes) alongside the usual workday using Evernote and (2) a supervised sensemaking phase using B&P. Both phases included a short training of the respective tool. The central part of the latter phase represented a fifteen minutes thinking-aloud usage of B&P at the workplace. The participants' task was to make sense out of their gathered bits: i.e. to first remember their informal learning experiences based on the bits with the help of time-based cues in the timeline and afterwards structure them semantically with the help of the circles in the organizing view. The second phase closed with a semi-structured interview asking about the participants' general attitude towards the prototype, its usability, performance and efficiency to support for remembering and sensemaking of their past experiences. The thinking-aloud and interview records were transcribed, paraphrased and categorized in an inductive, iterative process [6].

On average, the participants gathered eleven information bits over the two weeks collection phase (n=4; M=10.5; SD=2.52). The number of collected bits was rather low due to the voluntary participation, the workload the clinicians were facing during study time and problems with the unfamiliarity in using Evernote which influenced the study participant's motivation and/or ability in recording their own traces of informal learning experiences. The sensemaking phase worked out well except that the time needed for acquainting the participants with B&P was slightly underestimated. This resulted in more time spent on introducing B&P and a reduced sensemaking time for some.

In general all four participants were positive about the B&P tool, i.e. the study participants understood and appreciated the underlying idea and purpose. While for the study each participant had to choose a real learning need they currently faced, their motivation in using the tool appeared to vary. For the GP it was important to gather evidence for her revalidation, others saw it as an opportunity to pursue personal learning challenges or collecting information with the main purpose to share with others. For the latter, therefore, participants used Evernote mostly to gather todos for, e.g., reading and saving documents, Web links, etc. B&P was observed as being a responsive tool by three of the participants, although only easy to use by two of them.

**Timeline widget.** One participant's thinking aloud protocol clearly indicated that the timeline is working properly to prompt remembering informal learning experiences: 'That's the card sort photo, I remember that [...]. Just check I've remembered all the things I thought about last time.' However, although the idea was generally appreciated, most of the participants wanted extra cues to increase the chance to find and remember an informal learning experience, such as information on the involved persons or institutions, location, content and relation to other events. This confirms the former request raised in the paper prototyping session [10], hence proves the design decision to go for a widget-based framework easing the implementation of additional widgets with focus on required cues. Furthermore, we observed participants re-evaluating a bit after having remembered it: '[D]o I still need this? If I do need this, what am I going to do with it?'

That is, the person examines it more deeply than it was possible during busy working time and assesses it again. Another participant put it like that: ‘What you really need, is to be able to go back to something you’ve collected, go into it, read it and then [...] make your own notes.’ Consequently, most participants requested to directly manipulate the title, tags or importance of bits and attach a comment, reminder or deadline favorably accessible via context menus. As for the narrowing down of information, timeline anchors, searching functionality as well as filters for quicker browsing and comparing of bits was desired. These feature requests reflect the need for interactivity while sensemaking and support the respective design decision. In this way, the timeline reminds about past experiences and provides a view of things to work on at a glance much better than in an email client as one participant put it. Furthermore, the timeline indicates one’s own learning activity over a longer period of time and thus provides insight into the actual situation regarding revalidation as another participant noted.

**Organizing widget.** Overall, the interactive nature of sensemaking via the organizing and categorizing of bits into circles was valued by the healthcare professionals, particularly the opportunity to break topics down further was appreciated. In the same lines, the request to create circles within circles, to enable the creation of subcategories, was formulated as well. One participant was especially successful in her sensemaking effort as she managed to meaningfully group her collected informal learning experiences on diabetes treatment according to their influence on blood pressure: ‘[P]eople’s blood sugar levels have gone down through exercise as opposed to diet’. So, skim reading in the sensemaking phase resulted in a first internal representation of her learning, which lead to the creation of a corresponding external categorization in the circles ‘Results from diet’ and ‘Results from exercise’ and their overlap representing results from both. Having had more time for sensemaking, it is likely that the involved representations would have been enriched with more data, adapted and matured over time. This provides the first evidence that the approach of interactively organizing and categorizing bits into circles can support an individual in his sensemaking attempt properly. However, the majority of the participants did not believe this kind of sensemaking activity would influence their individual knowledge. As aforementioned, it turned out that people needed more time to actually grasp the content of collected bits, understand their relevance and re-evaluate them. The lack of time for this in the study design may have hindered their chance to engage in sensemaking. Finally, there was a strong demand of two participants to make the transition of pieces of information from the timeline to the organizing view permanent, i.e. the dragged bits disappear from the timeline. One participant put it this way: ‘[T]hey’re on the timeline [...] so I know now that I have to [...] read them[. T]hen once I’ve read them, put them into whichever category [...,] I would have thought to take that off the timeline means then that’s finished.’ This feature request is in line with underlying theoretical assumptions that once sense is made about informal learning experiences they are transferred from episodic into semantic memory, hence memorized in a distinct way. As a positive side-effect of removing processed bits from the timeline, it would get cleaned up and well-arranged.

## 4 Conclusion

This study contributed in progressing our understanding of sensemaking of informal learning experiences by validating a novel sensemaking interface called B&P in a realistic field setting with realistic learning needs. We found initial support for the effectiveness of the interface for supporting retrieval and foraging traces for informal learning experiences via a timeline and organizing view. Moreover, we found evidence for the need of a re-evaluation step as part of the sensemaking process. These results suggest that the current approach is valid and can serve as a basis for further developments.

With the B&P prototype we extend the HCI perspective of sensemaking [9][7] by considering sensemaking of informally collected traces for informal workplace learning. In our future work we will further refine this view by conducting a more sophisticated experimental design that also takes into account changes of internal knowledge representations. Furthermore, we plan to extend the tool to cover alternatives to the timeline and the organizing widget. Also, sharing features (among many others) are currently under development. With these steps we hope to obtain more insights into the problem of sensemaking of informal learning experiences and to develop a useful sensemaking tool in the end for healthcare professionals not only in the UK.

**Acknowledgement.** This work is supported by the Know-Center and the EU-IP Learning Layers. Learning Layers is funded by the European Commission within the 7th Framework Programme (Grant Agreement 318209). The Know-Center is funded by the Austrian COMET Program under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

## References

1. Butcher, K., Sumner, T.: Self-Directed Learning and the Sensemaking Paradox. Human-Computer Interaction 26(1), 123–159 (2011)
2. Dervin, B.: Sense-making theory and practice: an overview of user interests in knowledge seeking and use. Journal of Knowledge Management 2(2), 36–46 (1998)
3. Eraut, M.: Informal learning in the workplace. Studies in Continuing Education 26(2), 247–273 (2004)
4. Horng, D., Chau, P., Kittur, A., Hong, J.I., Faloutsos, C.: Apolo: Making Sense of Large Network Data by Combining Rich User Interaction and Machine Learning. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 167–176 (2011)
5. Kowald, D., Dennerlein, S., Theiler, D., Walk, S.: The Social Semantic Server. In: I-SEMANTICS (Posters & Demos), pp. 50–54 (2013)
6. Mayring, P.: Qualitative content analysis. Forum: Qualitative Social Research 1(2) (2000)
7. Pirolli, P., Card, S.: The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis. In: Proceedings of International Conference on Intelligence Analysis (2005)
8. Prilla, M., Degeling, M., Herrmann, T.: Collaborative reflection at work: supporting informal learning at a healthcare workplace. In: Proceedings of the 17th ACM International Conference on Supporting Group Work (2012)

9. Russell, D.M., Stefik, M.J., Pirolli, P., Card, S.K.: The cost structure of sensemaking. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 269–276 (1993)
10. Tomberg, V., Al-Smadi, M., Treasure-Jones, T., Ley, T.: A Sensemaking Interface for Doctors Learning at Work: A Co-Design Study Using a Paper Prototype. In: ECTEL Meets ECSCW 2013: Workshop on Collaborative Technologies for Working and Learning, pp. 54–58 (2013)

# Automatic Generation of SKOS Taxonomies for Generating Topic-Based User Interfaces in MOOCs

Carmen De Maio[2], Vincenzo Loia[2],
Giuseppina Rita Mangione[1], and Francesco Orciuoli[1]

[1] Dipartimento di Ingegneria dell'Informazione,
Ingegneria Elettrica e Matematica Applicata (DIEM), University of Salerno, Italy
{gmangione,forciuoli}@unisa.it
[2] Dipartimento di Informatica (DI), University of Salerno, Italy
{cdemaio,loia}@unisa.it

**Abstract.** The aim of the paper is to provide a framework for the automatic generation of topic-based user interfaces for video lectures in MOOCs. The proposed approach leverages on Fuzzy Formal Concept Analysis and Semantic Technologies, which allow the definition of solutions for supporting learners to navigate the fragments of one or more video lectures, by selecting the topics of interest. The exploitation of a Semantic Web vocabulary, namely SKOS, to model topics, and their relationships, enables the interconnection of different video lectures, also belonging to different MOOCs. The high interoperability allowed by Semantic Web technologies enables the integration of different and heterogeneous MOOC Platforms, as well as other Open Repositories. This aspect fosters the capability of learners to self-regulate and to enhance their learning paths in new forms of learning experiences based on exploration.

## 1 Introduction, Motivation and Related Works

Massive Online Open Courses (MOOCs) are becoming an important part of the education world. MOOCs are seen as a *disruptive innovation* [1] that starts from the OERs (Open Educational Resources) benefits and from the Social Learning main principles in order to create massive and professionalizing experiences. The main content delivery mechanism of a MOOC consists in a learning/teaching environment that provides streaming of video lectures, typically scheduled in weeks. These video lectures are aggregated in modules and courses, which are accessible for fixed time windows. In MOOCs, learners enjoy video lectures produced by the most prestigious Universities in the world (e.g. Berkley, MIT, Stanford, etc.) and take part in interactive experiences, which require an adequate level of autonomy.

Although recent researches suggest that video lectures increase student engagement and short-time videos provide learning outcomes comparable to traditional on-campus lectures [2], further investigations [8] provide negative results: only half the participants watches the majority of videos within a course. More in details, video lectures in MOOCs imply passive watching and mainly go without the use of interactive video

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 398–403, 2014.
© Springer International Publishing Switzerland 2014

tools. This model could be effective (in some cases) when the goal is to acquire elementary knowledge, but it does not support a good cognitive development. In massive environments, learners can check the correspondence between learning paths and goals, monitor their progress with respect to the scheduling of lessons and participate in assessment events and final exams, but they are not enabled to activate processes of permanent knowledge generation and navigation within extended cognitive spaces. In order to exploit the full potential of massive environments and video lectures and to overcome their limitations we propose an approach, based on Computational Intelligence and Semantic Technologies, to construct topic-based interfaces for browsing and accessing single parts of the aforementioned videos by considering the concepts each part deals with.

Thus, this paper proposes the definition of a methodology for video content analysis in order to extract conceptual information from video lectures and generate SKOS taxonomies representing the basis for the construction of the aforementioned topic-based user interfaces. Indeed, one of the main benefits of the proposed approach is providing learners with the chance to graphically and conceptually observe and browse video lectures and single fragments of video lectures. This capability allows to maximize the *generative cognitive processes* by minimizing *extraneous processing* and optimizing *essential processing* [10]. First of all, topic-based interfaces reduce, for learners, the percentage of irrelevant content with respect to their objectives and needs. Thus, it is possible to minimize the *extraneous cognitive load*. At the same time, the conceptualization of video lectures, realized by means of topic-based interfaces, enables learners to: i) connect their goals with key concepts of the lectures (personalization) by making explicit the conceptual structures; ii) connect the above mentioned key concepts with their prior knowledge (concretization) and, iii) access content by means of suitable semantic links (anchoring) to improve the whole learning experience.

## 2 Proposed Methodology

The proposed methodology (implemented as a software prototype in the context of this work) foresees a first phase of *Pre-processing* and a second phase (i.e., *Taxonomy Building*) where a workflow for building a taxonomy that conceptualizes the content of the considered video lectures is executed. Firstly, the *Pre-processing* phase consists in extracting texts from video lectures. This task can be accomplished in two different ways. The first one is to exploit transcripts (in different languages) provided by many MOOCs. The second one is to employ some techniques of *speech to text*. In our first software prototype, the *Java Speech*[1] library has been employed to perform the text extraction task. Secondly, the *Taxonomy Building* phase is realized by a workflow composed of several steps as shown in Fig. 1. This phase takes into account the textual content, extracted by means of the *Pre-processing* phase from video lectures of one or more MOOCs, and constructs a knowledge model represented by means of the Semantic Web technologies[2] (e.g., RDF, RDFS, etc.).

---

[1] `http://jsapi.sourceforge.net`
[2] `http://www.w3.org/2001/sw/`

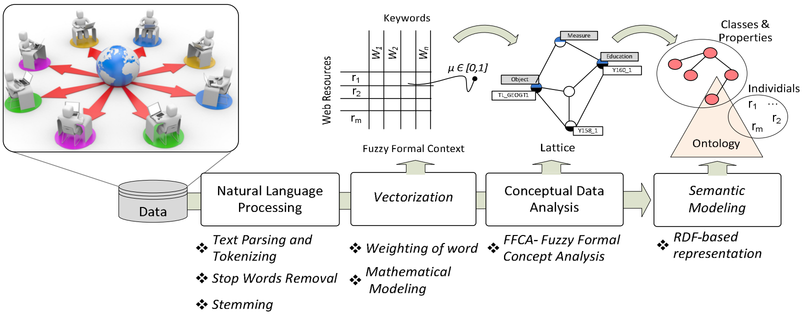



**Fig. 1.** Taxonomy Building

The workflow of the *Taxonomy Building* phase is essentially composed of the following steps: *Natural Language Processing*, *Vectorization*, *Concept Data Analysis* and *Semantic Modeling*. In particular, the last two steps (we are going to describe) are the most interesting.

*Concept Data Analysis* is aimed at extracting, from texts, concept hierarchies and relational schemas represented by means of a mathematical model, namely *fuzzy concept lattice*. Specifically, the fuzzy extension of *Formal Concept Analysis, i.e., Fuzzy FCA* [4], is applied. This step takes, as input, the vectors of words created in the previous steps (*Natural Language Processing* and *Vectorization*) and constructs the *fuzzy formal context*. Concept Data Analysis allows organizing fragments of video lectures (which are named objects) and features (which are named attributes) according to the shared meaning. Intuitively, we are interested in grouping together the greatest number of objects which share the same set of attributes, and vice-versa. More in details, it is possible to enable the representation of the relationships between objects and attributes in a given domain by means of the *fuzzy formal context*. Subsequently, starting from the *fuzzy formal context* it is possible to derive a *fuzzy concept lattice*, i.e., a mathematical structure, used to model knowledge, that is more informative than traditional tree-like conceptual structures [5]. The nodes of such structure are called *fuzzy formal concepts*.

Furthermore, *Semantic Modeling* is aimed at representing the extracted concepts (and their relationships) in a schema that is compliant with the Semantic Web technologies (i.e., SKOS and RDFS). The extracted *concept lattice* allows providing a machine-interpretable representation of knowledge in order to support integration, reuse, reasoning and querying. This process is crucial to enable the semantic annotation of video lectures.

### 2.1 Semantic Modeling: Taxonomy Extraction

*Semantic Modeling* is the process enabling the extraction of knowledge embedded in the video lectures and building the SKOS taxonomy that conceptualizes the above mentioned digital resources. SKOS *is a vocabulary supporting definition and use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading lists and taxonomies within the framework of the Semantic Web*[3].

[3] `http://www.w3.org/2004/02/skos/`

Firstly, each *fuzzy formal concept* of the *lattice* becomes an instance (individual) of the *skos:Concept* class [6]. The concept label (the most significant term associated with the current concept), selected in the Concept Data Analysis phase, is modeled with the *skos:prefLabel* property (related to a *skos:Concept* individual). All the other attributes, which are not inherited in the lattice structure, are considered related words for the current concept and are represented by using the *skos:hiddenLabel* property. In the example in Fig. 2, the concept *#C1* has the label ‘Operating System’, so it is enriched by means of the literal ‘Operating System’ as a *pref:Label*.

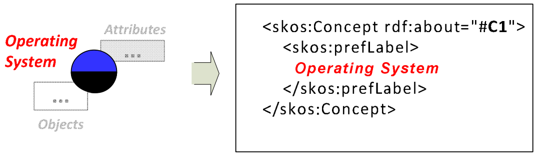



**Fig. 2.** Example of Concept Mapping

Once all the concepts (in the lattice) have been represented by means of *skos:Concept* individuals, a hierarchical mapping has to be performed. In this task, the hierarchical relationships are represented by using the *skos:broader* and *skos:narrower* (i.e., the inverse of *skos:broader*) properties which are used to connect, respectively, child-nodes with father-nodes and vice-versa. A wider discussion of how SKOS can be exploited in MOOC Platforms is considered in [12].

### 2.2 Inter-MOOCs Linking and Navigation

The main idea is to enable the interlinking of different MOOCs (also deployed on different technological platforms), which deal with the same domain or with overlapping domains. The definition of these links generates a topic-based interface to browse cross-MOOC contents. So, new MOOC experiences for learners come to life. In this work, a further methodology to link concepts belonging to different SKOS taxonomies (also coming from different MOOCs) is provided. The implementation of this methodology is based on the customization of Silk (Link Discovery Framework) [7], a tool designed to identify relations between entities belonging to different data sources. In particular, techniques for *words similarity* computation [11] have been implemented and injected into Silk that provides an XML-based declarative language that can be used to specify which types of links could be discovered between data sources as well as which conditions data items must fulfill in order to be interlinked. These conditions are based on different similarity metrics applied on entity properties which are addressed by using a path-based selector language [7].

### 2.3 Building the Facet-Based Interfaces

Recently, the semantic navigation of large classified datasets is mainly supported by the *Faceted Browsing* approach as emphasized in [9]. In this work, we enhanced the proposed methodology in order to generate Facet-based user interface from the previously defined SKOS taxonomies. Specifically, by means of Facet-based user interfaces, learners have the chance to drill down concepts (of SKOS taxonomies) and thus to consider multiple dimensions, namely *facets*, when they browse video lecture fragments. Each *facet* provides multiple restriction values, which allow learners to select and combine restriction values to constrain relevant items in the information space and navigate among the fragments of several video lectures. Definitely, the Facet-based approach provides a flexible mechanism to browse and search heterogeneous contents and enables exploratory experiences for learners in MOOCs.

## 3 Case Study and Evaluation

In order to provide an overview of the implemented system prototype, a case study has been executed and its results are reported in this section. The results (concepts, provided by the analysis described in Section 4, and their relevance for each analyzed video) of the extraction of conceptual information from two video lectures are shown in Fig. 3. In particular, the two video lectures provide content explaining the Semantic Web and its applications. The system extracts concepts like: *w3c, web, vagueness, reasoning, standard, berners-lee,* and so on.

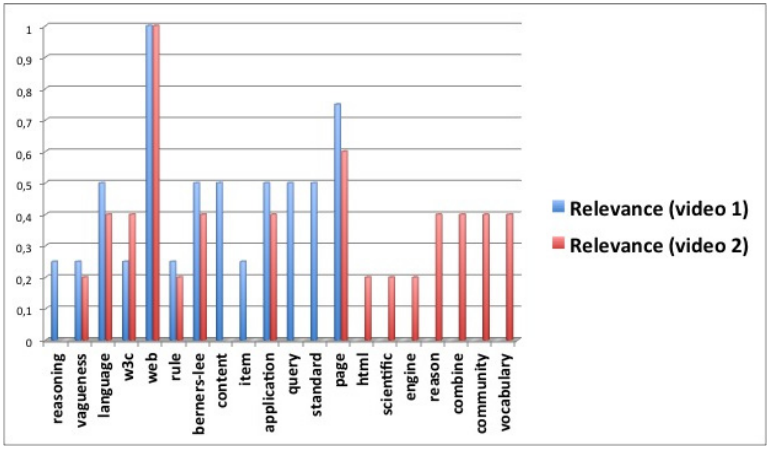



**Fig. 3.** Results of the early experimentation

## 4 Final Remarks

This paper proposes an approach, based on *Fuzzy Formal Concept Analysis*, to build semantic interfaces for browsing video lectures in MOOCs by using conceptual information organized in SKOS taxonomies. The aforementioned conceptual structures allow new forms of learning experiences in MOOCs by fostering topic-based exploration in the multimedia content space consisting in fragments of traditional video lectures of MOOCs. The proposed approach provides also additional relevant benefits. The first one is the interoperability. SKOS is compatible with the Semantic Web and Linked Open Data principles. This aspect improves searching and reusing processes for learning contents. The second one is the capability to support the interlinking of several MOOC environments. This aspect enables cross-MOOC browsing processes. Lastly, the results of the first experimentation are promising. For this reason, the authors will continue to improve the methodology and the system and will carry out further experimentations to evaluate also the quality of the learning processes.

## References

1. Baker, P., Bujak, K.R., DeMillo, R.: The evolving university: Disruptive change and institutional innovation. Procedia Computer Science 14, 330–335 (2012)
2. Glance, D.G., Forsey, M., Riley, M.: The pedagogical foundations of massive open online courses. First Monday 18(5) (2013)
3. Ramos, J.: Using TF-IDF to determine word relevance in document queries. In: First International Conference on Machine Learning (2003)
4. De Maio, C., Fenza, G., Loia, V., Senatore, S.: Hierarchical web resources retrieval by exploiting Fuzzy Formal Concept Analysis. Journal of Information Processing and Management (2011), ISSN 0306-4573
5. Zhou, B., Hui, S.C., Chang, K.: A formal concept analysis approach for Web Usage Mining. In: Shi, Z., He, Q. (eds.) Intelligent Information Processing II. IFIP, vol. 163, pp. 437–441. Springer, Heidelberg (2005)
6. SKOS Simple Knowledge Organization System Reference, World Wide Web Consortium (2009), `http://www.w3.org/TR/skos-reference/`
7. Bizer, C., Gaedke, M., Kobilarov, G., Volz, J.: Silk – A Link Discovery Framework for the Web of Data (2009)
8. Seaton, D.T., Rodenius, A.J., Coleman, C.A., Pritchard, D.E., Chuang, I.: Analysis of video use in edX courses. In: AIED 2013 Workshops Proceedings Volume, p. 57 (2013)
9. Tvarožek, M.: Exploratory search in the adaptive social semantic web. Information Sciences and Technologies Bulletin of the ACM Slovakia 3(1), 42–51 (2011)
10. Mayer, R.E.: Seeking a science of instruction. Instructional Science 38(2), 143–145 (2010)
11. Witten, I., Milne, D.: An effective, low-cost measure of semantic relatedness obtained from Wikipedia links. In: Proceeding of AAAI Workshop on Wikipedia and Artificial Intelligence: An Evolving Synergy. AAAI Press, Chicago (2008)
12. Miranda, S., Mangione, G.R., Orciuoli, F., Gaeta, M., Loia, V.: Automatic generation of assessment objects and Remedial Works for MOOCs. In: IEEE International Conference on Information Technology Based Higher Education and Training, ITHET 2013 (2013)

# Blueprint for Software Engineering in Technology Enhanced Learning Projects

Michael Derntl, István Koren, Petru Nicolaescu,
Dominik Renzel and Ralf Klamma

RWTH Aachen University
Advanced Community Information Systems (ACIS)
Lehrstuhl Informatik 5, Ahornstr. 55, 52056 Aachen, Germany
lastname@dbis.rwth-aachen.de

**Abstract.** Many projects in Technology Enhanced Learning (TEL) aim to develop novel approaches, models, and systems by field-testing new ideas with software prototypes. A major challenge is that project consortia need to establish a distributed, typically understaffed developer community that has to align its development efforts with needs from application partners and input from research partners. Tackling this challenge, this paper provides a blueprint for software engineering process and infrastructure, which was distilled from successful practices in recent TEL projects. We present a composition of freely available instruments that support open, distributed software engineering practices using continuous integration processes. The blueprint considers the full development cycle including requirements engineering, software architecture, issue tracking, build management and social aspects of developer community building in TEL projects. Some lessons learned are provided, particularly related to open source commitment, innovation as a social process, and the essential role of time. We aim to make software development in TEL fit for Horizon 2020 with processes and instruments that can be readily adopted for planning and executing future projects.

## 1 Introduction and Challenges

In interdisciplinary projects the negotiation of the main outputs and work processes is key to success. This is particularly true for *Technology Enhanced Learning (TEL)* projects, since these projects face the challenge of bridging a potentially huge gap between learning theories and the enabling technologies [1]. Most R&D projects funded by the TEL unit in FP6 and FP7—most notably specific targeted research projects and large-scale integrated projects[1]—have aimed to develop novel TEL approaches by evaluating and field-testing new ideas with software prototypes. Since the TEL community is small and tightly connected [2], it is a natural consequence that the consortia often follow approaches that have succeeded in previous or similar projects. These might include practices like rapid prototyping, continuous integration, agile development, and others [3]. As

[1] For an overview see http://www.learningfrontiers.eu/?q=project_space

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 404–409, 2014.
© Springer International Publishing Switzerland 2014

a complicating factor, collaborative research projects often face the challenge of an understaffed, distributed developer community that needs to align development efforts with needs from application partners and input from researchers. To resolve the forces in this area of tension, this paper presents a blueprint for software engineering in TEL research projects. The paper focuses on large-scale projects that require deployment and integration of scalable solutions and development processes based on available technologies, end-user involvement, and informed decision making.

The roots of the research presented in this paper lie in the technical leadership in two inherently different large-scale integrated projects in TEL, namely ROLE[2] and Layers[3]. While in ROLE the task was to develop and deploy a platform for responsive open learning environments supporting self-regulated learning (e.g., [4]), the challenge in Layers is to develop and deploy scalable, flexible and rapidly deployable infrastructures for informal learning [5] in two large SME clusters in the UK (health care) and Germany (construction). There is negligible overlap in functional requirements in these two projects, yet both exposed a considerable amount of common challenges and non-functional requirements related to the software architecture as well as the development and integration processes. Key architectural challenges are the need to make early architectural decisions, and to build the basis for flexible, customizable, traceable, and scalable solutions. Key development and integration challenges include the distributed developer community, the danger of a lack of stakeholder commitment, swift provision of software engineering infrastructure, and establishment of development practices.

Many of these challenges will be relevant to other large-scale R&D projects in TEL and beyond. It is understood that TEL innovation processes need alignment with research and practice that can build on previous findings [6]. In this spirit, to preserve and spread effective practices in TEL development we present in Section 2 successful solutions deployed first in ROLE and then adopted and refined in Layers in the form of a blueprint for software engineering in TEL to be reused in other projects that offer comparable scope and challenges. In the last section we provide lessons learned and conclude the paper.

## 2 Distilling the Blueprint

The blueprint tackling the challenges presented in the previous section integrates three core perspectives, as explained below and illustrated schematically in Fig. 1.

### 2.1 Stakeholder Perspective

Typical collaborative R&D projects involve goals and strategies of researchers, application partners, and developers. Put provocatively, researchers aim to publish high-impact papers; application partners want ready-to-use, custom tailored

[2] `http://role-project.org`
[3] `http://learning-layers.eu`

apps; and the developer force is distributed, understaffed, and members often pursue PhD research. It is obvious that this perspective will impose a divergent force upon most projects. To bundle the development capacity we propose to establish a *Developer Taskforce* that acts largely autonomously when negotiating and realizing short- to mid-term development objectives within the development roadmap framed by researchers, co-design partners, and project management. This development community may consist of project internal developer groups, involved student groups—e.g., in project-based learning (PBL) courses offered by project partners—and external groups like local OSS communities, established companies and emerging start-ups.

For making decisions with a potentially project-wide effect during the software engineering process, we propose to establish a governing body (in Layers the *Architecture Board*), in which relevant stakeholders are represented, and which makes binding decisions based on suggestions and input by project partners.

### 2.2 Continuous Integration Perspective

Integration shall impose a convergent force upon the project. It has several dimensions. First, there is the integration with the application partners, which is

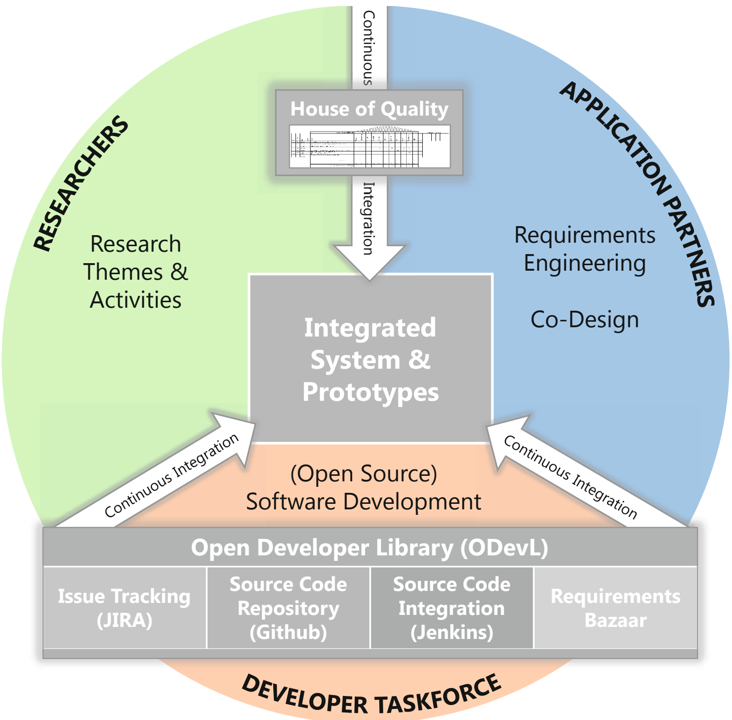



**Fig. 1.** Perspectives, activities and instruments in the software engineering blueprint

expressed in a systematic requirements engineering process. Especially in TEL projects, requirements engineering methodologies and tools must support an efficient communication and negotiation among the different stakeholder groups involved. Traditional requirements engineering techniques involving face-to-face interaction with end-users are resource-intensive [7], involving traveling costs and high preparation efforts. Such costs are justifiable in early project stages, when working processes are not yet established. However, in later project stages, such costs can be avoided by the use of effective and easy-to-use tools. As an open requirements platform we adopted Requirements Bazaar (see below). Second, there is the server-side integration, which facilitates the offering of unified services and tools developed in a distributed community. Third, there is the client-side integration that shall generate integrated end-user tools consuming common services. In addition there is continuous integration at the source code level (see below).

Integration efforts shall be driven mainly in *Co-Design Teams*, which subsume application partners, designers and developers as a permanent subgroup of the project working on a particular design challenge or design idea. They develop usage scenarios, wireframes, mockups, and first prototypes. Also they formulate demands for the further development of prototypes. Since there is often a wide variety of different design ideas in a project, it is of high importance to find the commonalities and key elements across these ideas. The assessment of alternative solutions shall be facilitated using the *House of Quality* (HoQ) approach [8], which pitches requirements against features offered by products to be adopted, developed and integrated. HoQ originated in Japan in 1972, and it was adopted during the 1980s by large U.S. companies such as Ford, Xerox and AT&T for their product development activities. It has proven to be a valuable instrument also in the Layers R&D context.

To make sure that co-design teamwork aligns with the overall project objectives, we propose to host project-wide *Integration Conferences* that focus on integrating the plethora of outputs from the co-design and development threads. This includes integration of theory, data, services, user interface, development processes and external tools.

### 2.3 Instrument Perspective

The key to let convergence win over divergence is the set of instruments deployed and offered to project partners to support continuous integration at all levels and in all areas where stakeholder interests meet. In particular we propose the *Open Developer Library (ODevL)* approach, which acts as a virtual hub offering tools and practices facilitating the Developer Taskforce:

– Source Code Repository: In a distributed software engineering process it is key for developers to have a reliable source code management in place. There are freely available solutions available. We suggest GitHub[4], since it is

[4] `http://github.com`

integrated with other tools and offers sophisticated support for a distributed development (e.g., full-fledged local repositories).
- Source Code Integration: Build-level code integration tools support and access source code management systems, e.g., to obtain code for regular integrated builds. The build process can be tailored to specific needs regarding the build triggers and the build process itself. In the blueprint we propose the open source solution Jenkins. Regular builds will offer timely notifications for those people in charge of integration.
- Issue Tracking: Issue trackers allow the management of software issues, including bugs, feature requests, etc. Most issue tracking software packages offer more than pure issue listings. The proposed system JIRA also offers project management tools (e.g., a Kanban board [9]), and integrated source code management tracking. The issue tracker is two-way integrated with the Requirements Bazaar.
- Requirements Bazaar: An open community toolkit that allows requirements and ideas generation and prioritization [10]. This is done in a bazaar-like metaphor, enabling negotiating around requirements using posting of artifacts, comments, and votes. This instrument provides a central requirements hub for all stakeholders in the project.

## 3 Conclusion and Lessons Learned

In this paper we have unfolded a blueprint for software development in large-scale R&D projects in TEL. We claimed that these projects often expose common challenges regarding the software architecture and the development process. We presented activities and instruments that allow approaching these challenges with proven tools and established practices in real-world developer communities.

The experiences about the software engineering process reported here are the synopsis of years of supporting TEL in creating and sustaining results within and beyond the scope of a funded research project. Some lessons learned include the insight that even with a good project plan, requirements and priorities of stakeholders change over time. These changes must be reflected in the process to avoid contracting the "not-invented-here" syndrome. Also it is advisable to make a strong commitment towards open source development, and to push the involvement of external developer communities to support the typically small project-internal developer force. A related issue which needs to be dealt with early is the OSS licensing models to be adopted. Also, time is an essential factor. It is paramount to start very early in providing the development infrastructure and grow continuously. In that sense, pre-configured development infrastructures like the blueprint presented in this paper, which can be rapidly deployed for a new project are a good choice. Last but not least, we emphasize the pivotal role of social factors in R&D projects; the success of deployed instruments will always depend on how well resistances can be resolved and how deeply stakeholders embrace the opportunity provided by such an infrastructure.

The actual resulting architecture was not explained in detail here since it will have limited general value outside of the Layers project. It is described in detail in [11]. With this paper we aimed to preserve previously and currently successful practice in a way that allows future TEL project consortia to plan and execute their software engineering processes, either as a dedicated work package, or by picking a subset of the instruments and activities that are tailored to their needs. We want to establish a culture of sharing and continued refinement of software engineering best practices in TEL across EU funded projects.

**Acknowledgments.** This research was co-funded by the European Commission through the FP7 Integrated Project "Learning Layers" (grant no. 318209).

## References

1. Noss, R.: 21st Century Learning for 21st Century Skills: What Does It Mean, and How Do We Do It? In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 3–5. Springer, Heidelberg (2012)
2. Derntl, M., Klamma, R.: The European TEL Projects Community from a Social Network Analysis Perspective. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 51–64. Springer, Heidelberg (2012)
3. Sommerville, I.: Software Engineering, 9th edn. Pearson, Boston (2011)
4. Carneiro, R., Lefrere, P., Steffens, K., Underwood, J. (eds.): Self-regulated Learning in Technology Enhanced Learning Environments: A European Perspective. SensePublishers, Rotterdam (2011)
5. Ley, T., et al.: Scaling informal learning: An integrative systems view on scaffolding at the workplace. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 484–489. Springer, Heidelberg (2013)
6. Beyond Prototypes consortium: Enabling innovation in technology-enhanced learning. Final project report (2013), http://is.gd/bpinnovtel
7. Law, E.L.-C., Chatterjee, A., Renzel, D., Klamma, R.: The Social Requirements Engineering (SRE) Approach to Developing a Large-Scale Personal Learning Environment Infrastructure. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 194–207. Springer, Heidelberg (2012)
8. Hauser, J.R., Clausing, D.: The House of Quality. Harvard Business Review 66(3), 63–73 (1988)
9. Anderson, D.J.: Kanban: Successful Evolutionary Change for Your Technology Business. Blue Hole Press (2010)
10. Renzel, D., Behrendt, M., Klamma, R., Jarke, M.: Requirements Bazaar: Social Requirements Engineering for Community-Driven Innovation. In: 21st IEEE International Requirements Engineering Conference (RE), pp. 326–327. IEEE (2013)
11. Derntl, M., Klamma, R., Koren, I., Kravcik, M., Nicolaescu, P., Renzel, D., et al.: Initial Architecture for Small-Scale Deployment. Learning Layers Deliverable D6.1(2013), http://is.gd/LayersD61

# Getting the Information You Need, When You Need It: A Context-aware Q&A System for Collaborative Learning

Irina Diaconita, Christoph Rensing, and Stephan Tittel

Multimedia Communications Lab
Technische Universität Darmstadt
Darmstadt, Germany
{FirstName.LastName}@kom.tu-darmstadt.de

**Abstract.** Keeping up with the rapid development in their fields is a problem for many employees nowadays, particularly when the new information is meant to be used for problem solving. Collaborative approaches such as forums and question and answering (Q&A) systems managed to only partly mitigate this problem in cases where time constraints and user mobility are of no crucial concern.

We present an enhancement to conventional Q&A applications which relies on user activity detection and question forwarding towards the available users. Our approach obtains a 90% precision in a garage experiment, proving itself as a viable solution for car technicians' time and mobility constraints.

**Keywords:** smartphone sensing, collaborative learning, context.

## 1 Introduction

Due to the rapid technical advancement in many fields, workers find it difficult to keep up with the required knowledge. Furthermore, in jobs that consist mainly of practical tasks and problem solving, one requires experience within similar problems more than theoretical knowledge.

Under these circumstances, colleague support is essential. However, traditional methods, like asking a colleague in person or over the phone pose a series of disadvantages. They generally limit employees to asking only the colleagues they have had previous contact with and carry the risk of interrupting colleagues during important tasks.

Therefore, question and answering (Q&A) systems offer a great opportunity for finding colleague support and sharing information. At the moment, however, they offer only limited help for certain fields where online communities are small and timely answers are crucial.

In this paper we present an approach meant to help car technicians by routing their questions to their available colleagues. Our concept could be used to implement similar systems for all kinds of technicians, from plumbers to airline technicians.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 410–415, 2014.
© Springer International Publishing Switzerland 2014

We base our concept on an user study carried out on 168 car technicians working in multiple garages in Germany. A third of the respondents reported often needing help, while the other two thirds found themselves only rarely needing external help. Results showed furthermore that 90% of the respondents are happy to help their colleagues. However, their willingness to help varied significantly based on their current activity: most respondents agreed to help during working hours, but fewer conceded to be interrupted from ongoing tasks and even fewer were willing to use their breaks for helping colleagues. Detailed results of the study can be found in [1].

Our paper is structured as follows: in Section 2 we look into the related approaches. Section 3 introduces our concept and Section 4 presents its implementation. We continue with the evaluation results in a real-life setting in Section 5. Finally, we present our conclusions and outlook in Section 6.

## 2 Related Work

Question and answering systems have been a quickly developing topic lately, both in research and in industry (Stack Overflow, Yahoo! Answers, etc.). One popular direction includes the analysis of the factors that influence the success of specific Q&A system. White et al [2] study the effects of community size and contact rate in social Q&A systems, while Mamykina et al [3] study the factors and characteristics that led to Stack Overflow's indisputable achievement of having about 92% of their questions answered in an average of eleven minutes.

A major direction, to which our work subscribes, is Q&A systems improvement. Shah et al [4] attempt to create an automated Q&A system from the existing knowledge base. Other solutions focus on capturing and reusing the expertise [5].

The other direction of our work concerns mobile sensing and user activity detection. Applications which make use of environment sounds include music and genre classification [6], event detection [7], cough detection [8] and phone position classification [9].

Our previous work [10] proposed a concept for a context-aware Q&A system using GPS and accelerometer data, as well as information from the smartphone's applications to detect user's activity. In contrast to that, our current approach employs audio recordings, fundaments its concept in a user study and has a real-world evaluation to prove our results. To the best of our knowledge, there are no other approaches for Q&A systems forwarding questions depending on user's current activity.

## 3 Concept

The backbone of our system is a Q&A infrastructure, that allows users to send and receive questions, as well as receive answers via multiple channels: smartphone app, web form, email, SMS. We thus comply with the technicians' mobility and different availabilities of the communication channels in their daily activities.

The car technicians' disponibility to help their colleagues varies across their daily activities. Therefore, each technician has to decide for himself if he is available or interruptible during each type of activity. We distinguish between availability as the state of a user that is willing and able to help during a certain activity and interruptibility, as the state of a user that is principally unwilling to help during his current activity, but could take a break for an urgent open question.

Our goal is to detect the technicians' current activities and, based on them, their availability and the interruptibility. Our centralized server would keep track of all the employees' availabilities and, when a new question arrives, it would select a few available technicians to whom it would forward the question. One of the precautions would be to prefer users that received or answered less questions during that day, so that the system does not burden the same few experts while leaving all the other colleagues out of the loop.

Another reason for which we need to know the user's current activity is so that we can use an appropriate communication channel (that he has assigned himself for that particular activity). For instance, a user cannot read and type into an app while driving, but can listen to a voice mail and record an answer to the question. Table 1 shows a sample mapping between the users' activities, their availabilities and interruptibilities, as well as the preferred communication channels.

**Table 1.** Car technicians' main activities

| Activity | Availability | Interruptibility | Communication channel |
|---|---|---|---|
| Car repairment in the garage | false | true | app/SMS |
| Car repairment on the street | false | false | - |
| Other work in the garage | true | true | app/SMS |
| Driving for work purposes | true | true | voicemail |
| Meeting | false | false | - |
| Break | false | true | app/SMS |
| Commute to/from work | true | true | voicemail |
| Other activity | true | true | app/SMS |

In order to detect the technicians' activities, we use their smartphones to record short audio samples. These are sent to our server, where they are classified. Thus, the samples undergo windowing, feature extraction and then classification.

## 4 Implementation

The Q&A system itself uses Microsoft SharePoint as a server, the questions being represented as threads in the message board. Questions can be submitted and answered both via a web interface and a web client. For managing notifications we use Redis, a pub/sub system.

Besides the main Android app, we also use a background service for gathering the audio recordings. This uses a wakelock to record sound samples at a given time and with a given duration.

On the server side, the sampled audio signal is first split into equally sized windows. Our implementation supports rectangular, Hamming and Hanning windows. Afterwards we apply a fast Fourier transform to convert each window to the frequency domain. In the next step we extract the feature vector for each window. We use four types of features: MFCC (Mel-Frequency Cepstrum Coefficients), Delta MFCC, Powerspectrum and Band Energy. As classifiers we use K-Nearest Neighbours (KNN), Decision Trees (DT), Random Forest (RF), Gaussian Naive Bayes (GNB) and Gaussian Mixture Model (GMM). The results of the classification concern the user's current activity, which is then mapped according to his previous personal settings to a availability and interruptibility status.

## 5 Evaluation of Activity Detection

To validate our approach to detect the activity, we carried out an experiment during which two car technicians used our application during their whole working time for two weeks. They both used Samsung Galaxy S4 phones, which could support a 44.1 kHz audio sampling rate. In order to collect our ground truth, the two technicians filled in separate daily forms, where they would specify all their activities and their corresponding time slots.

### 5.1 Comparison between Classifiers

In what follows, we will present and compare our classification results, using as measures the precision, recall and $F_1$ score. Figure 1 shows the performance of all combinations of feature vectors and classifiers that we used. As expected, MFCC and DMFCC yield the best results as far as feature vectors are concerned, as they are derived from the spectrum of the spectrum.

Random Forest, K-Nearest Neighbours and Decision Trees are the best performing classifiers. The clustered nature of the data, together with the overlapping of sounds and situations, like in the case of repairing a car outside versus in a garage, leads to Random Forest and K-Nearest Neighbours outperforming Decision Trees.

### 5.2 The Activity Classification in Detail

Next, we will analyze in depth the best classification results, which were provided by Random Forest Classifier together with MFCC features. Table 2 shows the precision, recall, $F_1$ score and support for each of the situations and for the overall classification. The overall precision of 90% could be achieved thanks to the distinct acoustic patterns of the situations. This also led to a clustering of data, which caused Random Forest to outperform other classifiers.

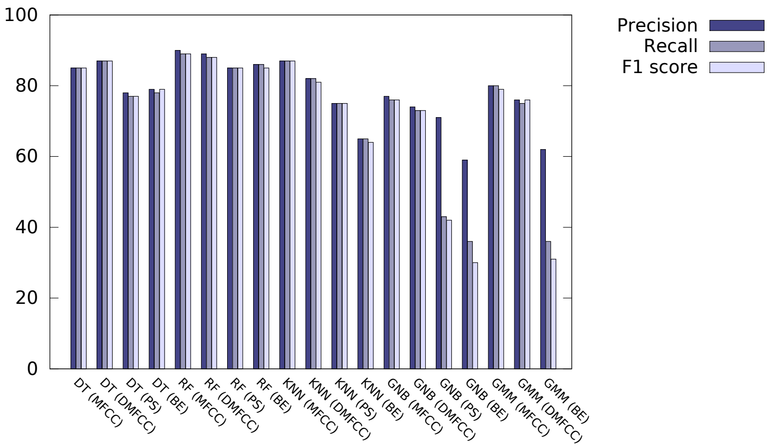



**Fig. 1.** Comparison of the performances of all the combinations of feature types and classifiers

Two apparently close situations, repairing a car in the garage and repairing it outside of the garage are distinguished remarkably well too. While a technician might perform similar tasks when repairing a car, regardless of the location, when a car is repaired in the garage there will often be more then one technician working there. Thus, most of the recordings taking place inside the garage contain the overlapping sounds of more cars being repaired, while the recordings of cars being repaired on the street correspond to sounds of only one car being repaired.

**Table 2.** Evaluation of the classification results of Random Forest and MFCC

| Class | Precision | Recall | $F_1$ score | Support |
|---|---|---|---|---|
| Break in the garage | 0.95 | 0.85 | 0.90 | 212 |
| Meeting | 0.87 | 0.95 | 0.91 | 724 |
| Repairment in the garage | 0.86 | 0.90 | 0.88 | 298 |
| Repairment outside of the garage | 0.95 | 0.84 | 0.89 | 390 |
| Other work in the garage | 0.8 | 0.21 | 0.33 | 19 |
| Overall | 0.90 | 0.89 | 0.89 | 1643 |

## 6 Conclusions

We have presented our approach for a context-aware Q&A system. Our concept has been motivated by a thorough user study. We have evaluated our activity detection approach with real-life data gathered in garages and and achieved a 90% precision.

Our next step would be to carry out an extensive user study in order to evaluate the system as a whole and not only the accuracy of the activity detection. Another important future improvement will be the implementation of the classifier on the phone, in order to alleviate privacy and user data protection concerns.

**Acknowledgment.** This work is supported by funds from the German Federal Ministry of Education and Research under the mark 01PF10005B and from the European Social Fund of the European Union (ESF). The responsibility for the contents of this publication lies with the authors.

## References

1. Rensing, C., Diaconita, I.: A Q&A System Considering Employees? Willingness to Help Colleagues and to Look for Help in Different Workplace-related Situations. In: Proceedings of the IEEE International Conference on Advanced Learning Technologies (2014)
2. White, R.W., Richardson, M., Liu, Y.: Effects of Community Size and Contact Rate in Synchronous Social Q&A. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2011)
3. Mamykina, L., Manoim, B., Mittal, M., Hripcsak, G., Hartmann, B.: Design Lessons from the Fastest Q&A Site in the West. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2011)
4. Shah, C., Oh, J.S., Oh, S.: Exploring Characteristics and Effects of User Participation in Online Social Q&A Sites. First Monday (2008)
5. Budzik, J., Hammond, K.: Q&A: A system for the Capture, Organization and Reuse of Expertise. In: Proocedings of the Annual Meeting - American Society for Information Science (1999)
6. Wang, A.: An Industrial Strength Audio Search Algorithm. In: Proceedings of the 4th Symposium Conference on Music Information Retrieval (2003)
7. Rossi, M., Feese, S., Amft, O., Braune, N., Martis, S., Troster, G.: AmbientSense: A Real-Time Ambient Sound Recognition System for Smartphones. In: Proceedings of the International Conference on Pervasive Computing and Communications Workshops (2013)
8. Larson, E.C., Lee, T., Liu, S., Rosenfeld, M., Patel, S.N.: Accurate and Privacy Preserving Cough Sensing Using a Low-cost Microphone. In: Proceedings of the 13th International Conference on Ubiquitous Computing (2011)
9. Diaconita, I., Reinhardt, A., Englert, F., Christin, D., Steinmetz, R.: Do You Hear What I Hear? Using Acoustic Probing to Detect Smartphone Locations. In: Proceedings of the International Conference on Pervasive Computing and Communications Workshops (2014)
10. Diaconita, I., Rensing, C., Tittel, S.: Context-aware question and answering for community-based learning. In: Die 11. E-Learning Fachtagung Informatik der Gesellschaft für Informatik e.V. (2013)

# A Training Framework for Adaptive Educational Hypermedia Authoring Tools

Conor Gaffney, Thanos Staikopoulos, Ian O'Keeffe,
Owen Conlan, and Vincent Wade

KDEG, School of Computer Science & Statistics, Trinity College, Dublin
{cogaffne,Ian.OKeeffe,Athanasios.Staikopoulos,
Owen.Conlan,Vincent.Wade}@sccs.tcd.ie

**Abstract.** Authoring tools are widely accepted to be one of the most promising approaches to increase the widespread uptake of Adaptive Educational Hypermedia (AEH). AEH authoring tools have however not been as successfully adopted as was initially expected. This paper presents a clearly defined set of comprehensive training steps devised to increase the accessibility of AEH authoring tools. The objective of these training steps is to insure AEH authoring tool users have a clear understanding of how personalization will affect and benefit their courses. The paper includes an evaluation of the presented training steps.

**Keywords:** personalization, authoring tool, composition, training, education.

## 1 Introduction

Personalized eLearning [1] allows systems to adapt to the individual needs of the learner and improve the educational effectiveness of each learner's experience. Adaptive educational systems can take into consideration many aspects of a learner's requirements [2]. Personalized eLearning has been shown in many studies to be highly effective at increasing students' performances [3] and has been identified as an essential element of next generation eLearning systems [4]. A key challenge in mainstreaming adaptive eLearning is the difficulty and complexity in design and implementation associated with authoring [5,6].

A number of authoring tools have been designed and implemented so as to improve the accessibility of Adaptive Educational Hypermedia (AEH) composition [5, 7]. These tools and other Adaptive Hypermedia (AH) systems have been examined and evaluated in great detail [8]. The result of this scrutinization is that the design of AEH authoring tools are constantly being advanced and great efforts have been made to improve a users experience. While it is well documented that authoring AEH is difficult and complex [5] uptake of authoring tools has not advanced in line with their improved design. The reason for slow uptake may not lie in the design of the tools themselves but the training provided to users in their operation.

As presented in the results of the authors previous research [7] providing training related to AEH authoring is difficult. A study completed with the ACTSim Authoring

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 416–421, 2014.
© Springer International Publishing Switzerland 2014

Tool indicated that users found it challenging to understand the concepts of inclusion/exclusion related to adaptive navigation: *"users have a difficulty when they attempt to relate how their adapted model will be realised as an end personalised simulation"*. Foss and Cristea [5] include this concept in a set of authoring imperatives and state that AEH systems should accommodate a *"shallow learning curve"*. To accomplish this, users need to be adequately trained in both the technical application required for tool operation and the Adaptive Hypermedia theory that is needed to compose AEH coursers.

This paper presents a framework designed to structure training for AEH authoring tools. This innovative approach is flexible and easily extensible to include additional aspects of required training. The framework is based on existing design theory that is combined with state of the art of AEH authoring processes. This paper initially introduces a training framework for AEH Authoring Tools in the following section. This is followed by the presentation of a case study and evaluation with the final section describing the conclusions of this paper.

## 2 A Training Framework for AEH Authoring Tools

The objective of this training framework is to define a clear set of steps that can be used to improve the training delivered for the operation of AEH Authoring Tools. The framework is based on an instructional design framework for authentic learning [9] that incorporates situational examples. The framework incorporates typical design processes employed by some authoring tools that are based on adaptive hypermedia multi-model architectures. The following sections present the instructional design framework theory, AEH multi-model architecture authoring processes and an outline of the derived framework for AEH authoring tool training.

**Instructional Design Framework Theory:** Training for AEH Authoring Tools is typically delivered by an instructor or through automated online instruction. In either case the training is typically completed in an office environment with the aid of a computer. The training is also typically completed in the same context in which the course development will be completed. Herrington and Oliver [9] define a three parts in an instructional design framework for situated learning environments. The first part is to identify critical characteristics of the situated learning. The second part is to operationalize the critical characteristics within a multimedia program. The third part is to investigate the users perception of their experience. The framework includes the provision of including authentic contexts and examples.

**AEH Multi-model Architecture Authoring Processes:** AEH architectures are typically based on a multi-model approach [10] whereby user models, content/domain models and adaptivity/narrative models are all defined separately within a systems architecture. There are two main approaches to capturing the multi-model approach in AEH Authoring Tools. AEH Authoring Tools either require users to define multiple models to capture all the system requirements or a single model with multiple layers with each layer relating to an architectural component. For example, within the MOT authoring tool, users create content separately to an adaptivity model. The GRAPPLE

Authoring Toolkit (GAT) [10] employs a similar approach. The approach employed in the ACCT and ACTSim authoring tools differs to the MOT and GAT and requires users to initially define the content models with adaptivity control overlaid. The important aspect of these authoring processes with respect to training is that there is a clear separation of concerns for the SMEs when they are learning to use an AEH Authoring Tool.

**The AEH Authoring Tool Training Framework:** There are different aspects of an adaptive course that must be considered by an SME author. As evident in the GAT, ACCT and ACTSim authoring processes these can generally be defined as 1) the content and learning environment aspects and 2) the adaptivity controls. Content components typically include modeled concepts related to training material. Learning environment aspects refers to educational interventions, underlying operational models such as those found in eLearning simulations or immersive elements such as those employed in serious games. The type of AEH educational system influences the design of the content and learning environment aspects. The realization of these components can differ depending on the design of the Authoring Tools and will be unique to each system. It is proposed within this research that each of these architectural components are taught separately within AEH Authoring Tool training as reflected in AEH Multi-Model Architecture Authoring Processes. In the case of models, controls or environmental aspects being overlaid upon one another the *order* in which these architectural components are taught becomes important. If there is a clear separation of models within the authoring tool the order in which the authoring components are taught is less important. The final section of the training is typically related to the adaptivity controls of the AEH Authoring Tool as their explanation must be contextualized with other authoring components.

AEH Authoring Tool training is explained by following the Instructional Design Framework Theory previously outlined. This framework is applied across the explanation for each authoring component and the adaptivity controls. The training follows the three steps described in the theory and includes an introduction and examples where appropriate. The AEH Authoring Tool training steps are as follows:

<u>Introduction:</u> The first step of training to be presented to users is an introduction. It includes an *overview* of the training and outlines the *benefits of personalization.*

<u>Authoring Components:</u> These critical characteristics of authoring are each presented to users separately and, where appropriate, in sequence. For example, the ACTSim would require an explanation of the simulation-operating model (Dialogue Model) and learning interventions. The training presentation should include the following elements for each authoring component: *explanation of the authoring component; authoring tool architecture; example of the authoring component; a multimedia program presentation;* and *confirmation of user understanding.*

<u>Adaptivity Controls:</u> The final step of the training framework is to present the adaptive controls that will be employed in the AEH Authoring Tools. There are several different approaches employed to capture personalization in an authoring tool. For example, ACTSim employs tagging while the GAT is rule based. The training presentation should include the following: *explanation of adaptivity controls*;

*adaptivity controls affect; example of the adaptivity controls; a multimedia program demonstrating the adaptivity controls;* and *confirmation of user understanding*.

The following section presents a case study that outlines the AMAS Authoring Tool, a new and unique system designed to allow SMEs to compose adaptive activity based online courses. While the AMAS authoring tool was designed to be accessible to non-technical users it required a clear concise training framework to inform users of its operation while also providing a relevant introduction to Adaptive Hypermedia.

## 3 Case Study: The AMAS Authoring Tool

The AMAS Authoring Tool is a next generation AEH composition tool developed as part of the Adaptive Media and Services for Dynamic Personalisation and Contextualisation (AMAS) project [11]. The authoring tool allows non-technical SMEs to compose adaptive activity based online courses. The AMAS Authoring Tool was developed so as to integrate into pre-existing the AMASE architecture [12]. SMEs create adaptive activity based courses by defining a workflow graph that captures the course components with connected nodes. The authoring tool employs three types of authoring component, in a nested structure, to describe the content and learning environment aspects of the course: Activities, Tasks and Topics. Tags are used in the workflow by SMEs to describe the adaptivity. All of the components are captured and configured in a single model so as to reduce the complexity involved in authoring. This approach results in authoring components being layered on top of one another. Multiple models are required by the AMASE architecture for publishing and are extracted from the single model that is created by SMEs. A more detailed description of the AMAS Authoring Tool is presented in previous publications [13].

### 3.1 AMAS Authoring Tool Training

Training for the AMAS Authoring Tool was defined by following the AEH Authoring Tool Framework previously presented. Microsoft PowerPoint was employed to script and deliver the training. Microsoft PowerPoint was selected as it allows multimedia presentations to be easily implemented with the use of animation. Text, diagrams and images were also included to create a complete training package.

A total of twenty slides were employed to deliver the training that consisted of three main sections: the introduction; authoring components; and adaptivity controls. The training started with an introduction that outlined the benefits of activity-based training and the advantages of adapting these learning environments to learners. The training then presented an explanation of each authoring component employed in the AMAS Authoring Tool: Activities, Tasks and Topics. The examples employed to demonstrate the nesting were abstract in nature. The final slide describing the authoring components was a multimedia animation. It displayed an example course that included an animation to highlight the operation of the workflow. SMEs were asked to confirm that they understood the training related to the authoring components. Following the presentation of the multimedia animation SMEs were presented with the adaptivity control training. Adaptivity control training included an explanation of the Tags, a description of the affect the Tags will have on the course, an example of Tags and a multimedia animation demonstrating how Tags affects different users completing a course. The final part of the training included a demonstration of the AMAS Authoring Tool functionality. The demonstration displayed how each of the authoring components and adaptivity controls were added and configured in course workflow. To evaluate the training presented SMEs were requested to complete a questionnaire. Details of the user trial are presented in the following section.

### 3.2 Evaluation

The objective of the evaluation was to examine the effectiveness of the framework presented in section 2. In order to assess the framework a trial based user study was employed. An instructor delivered the training to each SME in their work environment. Once the instructor had delivered the training, the SMEs were asked to complete a questionnaire designed to gather both quantitative and qualitative data. The questionnaire was divided into three parts, each examining one of the three different sections of training. Quantitative data was gathered via Likert scales and qualitative data gathered by SME text input. A total of 8 authentic SMEs (teachers, trainers, lecturers and pedagogical designers) completed the evaluation after being presented with the AMAS Authoring Tool training.

The results of the evaluation are divided into three parts, one for each of the training sections presented to the SMEs.

**Introduction:** The results of the Introduction training section were mixed. All of the participants agreed that the training regarding the benefits of *adaptivity* was acceptable. However, three of the users indicated that the benefits of *online activity learning* were presented inadequately. A possible reason for this poor result is that the concept of *'online learning'* was not separated within the training from the concept of *'activity based learning'*. Although rudimentary concepts within eLearning, these would have been foreign concepts to some of the SMEs. The benefits related to *adaptivity* was explained and presented without interference of other concepts.

**Authoring Components:** The results related to the authoring components were positive. All of the participants indicated that the training related to the Activities, Tasks and Topics was acceptable and understood. Similarly, the examples employed were identified as being clear and the workflow (structuring of the components) was understood by all participants. The only exception to these results was that one user did not agree that the nesting of the authoring components was adequately presented.

**Adaptivity Controls:** The results of the evaluation related the adaptivity controls were the most positive of the three sections of training. All of the participants agreed that the training that explained the Tags, the examples and the animation presented were clear and concise.

While participants in the study were given the opportunity to include text based comments at the end of each section there was very little feedback gathered. The qualitative data that was gathered was generally positive. One participant commented the training *"was clear"* while another commented that it was *"excellent"* and that the *"examples situate it for learner, mapping examples to their reference of knowledge"*. The only somewhat negative comment related to the nesting of the authoring components. The participant commented that it *"might help to demonstrate Activity and Task and Topic in a flowchart form"*.

## 4 Conclusion

This paper has presented a new and unique framework developed by blending instructional design theory and the state of the art of AEH authoring tool authoring process design. The paper also presented the AMAS Authoring Tool, a new and unique authoring system that allows non-technical SMEs to rapidly develop adaptive activity based courses. The paper outlined the AMAS Authoring Tool training that was developed with the aid of the newly developed framework. Finally, the paper included an evaluation of the training that validated its design. Authentic real world subject matter experts were employed for the evaluation. The training is validated further by the fact that each of the SMEs went on to use the AMAS Authoring Tool to successfully compose the outline of adaptive activity based courses in their domain.

**Acknowledgments.** This work was funded by Science Foundation Ireland via grant 08/IN.1/I2103.

## References

1. Brusilovsky, P.: Adaptive educational hypermedia. In: 10th International PEG Conference, Tampere, Finland, June 8-12, pp. 8–12. Kluwer Academic Publishers (2001)
2. Wade, V.: Challenges for Multi Dimensional Personalised Web. Invited Keynote. In: International Conference UMAP, Trento, Italy (June 2009)
3. Sampson, D., Karagiannidis, C.: Personalised learning: Educational, technical and standardisation perspective. In: Interactive Educational Multimedia, vol. 4. AACE (2002)
4. Brusilovsky, P.: KnowledgeTree: A distributed architecture for adaptive e-learning. In: The 13th International World Wide Web Conference (WWW). ACM (2004)
5. Foss, J.G., Cristea, A.I.: The next generation Authoring Adaptive Hypermedia. In: Conference on Hypertext and Hypermedia, pp. 83–92. ACM (2010)
6. Nurjanah, D., Davis, H.: Improving the Workspce Awreness of Authors in Asynchronous Collaborative Authoring of Learning Designs. In: EdMedia AACE (2012)
7. Gaffney, C.: Authoring Adaptive Soft Skill Simulations. PhD. Computer Science. Dublin, Trinity College Dublin (2013)
8. Gena, C.: Methods and techniques for the evaluation of user-adaptive systems. Knowl. Eng. Rev. 20(1), 1–37 (2005)
9. Herrington, J., Oliver, R.: An instructional design framework for authentic learning environments. Education Technology Research and Development 48(3), 23–48 (2002)
10. Conlan, O., Wade, V.P., Bruen, C., Gargan, M.: Multi-model, metadata driven approach to adaptive hypermedia services for personalized elearning. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) AH 2002. LNCS, vol. 2347, pp. 100–111. Springer, Heidelberg (2002)
11. Ploum, E.L.M.: Authoring of Adaptation in the GRAPPLE Project. Technische Universiteit Eindhoven (2009)
12. OKeeffe, I., Staikopoulos, A., Rafter, R., Walsh, E., Yousuf, B., Conlan, O., Wade, V.: Personalized Activity Based eLearning. In: 12th International Conference on Knowledge Management and Knowledge Technologies (2012)
13. Gaffney, C., Conlan, O., Wade, V.: The AMAS Authoring Tool 2.0: A UX Evaluation. Paper Accepted at the Hypertext Conference, Chile, (2014)

# Analysis of Concept Sequencing in Student Drafts

Samuel González-López and Aurelio López-López

National Institute of Astrophysics, Optics and Electronics, Tonantzintla, Puebla, México
{sgonzalez,allopez}@inaoep.mx

**Abstract.** The proposal draft is the first step to achieve a university degree by students in many educational institutions. This proposal is transformed into a thesis after several revisions by an academic adviser. Each proposal must comply with requirements of institutional guidelines. In this paper we present an analyzer to identify the sequence of concepts within the proposal drafts in Computer Science and Information Technologies, with the goal of aiding students to improve their drafting. We propose four methods integrated into our analyzer, which were designed considering the transitions of grammar constituents (subject and object) in the sections of *Problem Statement*, *Justification* and *Conclusions.* We developed the methods from corpora analysis, validating them further. Moreover, we found that the sequences of concepts correspond to what is expected in the different sections.

**Keywords:** Technical writing, conceptual sequences, student drafts.

## 1 Introduction

In educational institutions, academic programs or courses often conclude with the writing of a thesis or research proposal, undertaken by students. These documents must comply with the drafting characteristics established by institutional guidelines, such as format, orthography, structure, legibility, or argumentation. Moreover, they have to satisfy the appropriate structural features of each section of a thesis. The usual process that the students follow is to write a first draft and then improve the document with the iterated reviews and recommendations of an advisor. Our efforts aim to aid students in preparing their proposal drafts by examining the sequence of concepts occurring in specific sections of their draft using an analyzer. We explore the relationship between paragraphs in Justification, Problem Statement and Conclusion sections, this relation is what we call *conceptual sequence*, that carries an implicit coherence. The key contribution of this work is the development of four methods to identify the behaviors in each selected section, based on a previously proposed technique for assessing local coherence.

We propose four different schemas to evaluate the conceptual sequences: First Paragraph of Reference, Evaluation of Nearest Neighbors, Cascade Evaluation, and Auto-evaluation of Paragraphs (section 3). These methods were incorporated into an analyzer. We explain the rationale of the four schemas and how they were formulated from corpora.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 422–427, 2014.
© Springer International Publishing Switzerland 2014

## 2 Background

A document with an appropriate structure presents a clear flow of topics through their paragraphs. For instance, in the paradigm of a five paragraph essay [1], the introduction and conclusion share the same main topic. The topic is the theme or subject matter of the essay. The remaining paragraphs in this approach named "body paragraphs" contain details of the essay argumentation and are linked via the main topic. This approach is similar to a proposal drafts, e.g. in *Conclusions* section, where the paragraphs are connected by the same topic, hopefully that of the problem statement, and are expected to include support of the results.

The connection between paragraphs involves the interconnection of each of the sentences through its grammar constituents as subject and object. These constituents are observed as a pattern and allow to correctly interpret the information in the text. The sequence presented in the pattern can characterize specific types of discourses and therefore contribute to the assessment of the quality of the generated text [2]. In our work, e.g. in the *Problem Statement* section, the set of sentences that integrate each paragraph are interconnected by the same central topic. These connections provides an adequate sense of what the student seeks to address in the proposal. Under this approach, our research takes as a working unit the sentence.

In [3], the authors present a related study with semantic flow of essays of students learning English, and argue that the paragraphs of the essays have an internal flow, i.e. each paragraph connects with the adjacent paragraph. If paragraphs are connected in ascending order, the essay is suitable from a semantic approach (see Figure 1). They took the measure of distance between vectors representing paragraphs to determine their semantic proximity. This proximity depended on the topic of each paragraph. However, the topic flow in the essays was small, but present. They concluded that is possible to obtain a better semantic flow on collections of essays where the quality differences are more significant.

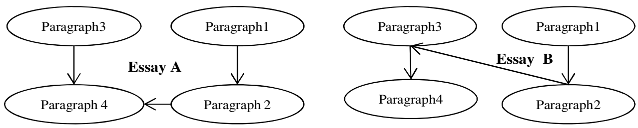



**Fig. 1.** Essay A with a low grade and the Essay B with a high grade of semantic flow

Our work takes advantage of a tool based on the Entity Grid (EGrid) technique to represent discourse, which was proposed to evaluate coherence in text [4]. The technique generates a representation constructed as a two-dimensional array that captures the distribution of entities in discourse across sentences, where rows correspond to the entities of discourse and the columns represent the sentences. The cells can have values such as subject (S), object (O), or neither (X). The main idea of this representation is that if similar object and subject are present in contiguous sentences, the assessed coherence is stronger. For example, given the sentences (main entities bracketed):

S1: [Microsoft]O is accused of trying to forcefully buy into [markets]X where [its own products]S are not competitive enough to unseat [established brands]O.

S2: [Microsoft]S claims [its tactics]S are commonplace and good economically.

we observe how the entity *Microsoft* appears initially as O in S1, then as Subject in S2, revealing certain coherence. The EGrid technique generates a model which is built from a specific corpus and such model is further used to evaluate new texts. The authors used 3,991 samples for training and 4,143 were used for testing on the topics of earthquakes and accidents. In our work, we identify the subject and object of sentences with the goal of detecting the sequence of concepts, instead of local coherence. But our proposal uses paragraphs to train the model that will assess the concepts of the rest of the section.

## 3 Analyzer Model

A proposal draft consists of a series of elements (sections) that have been established by books on research methodology and guidelines. In this work, we selected the sections of *Problem Statement, Justification* and *Conclusions* because they are longer texts of multiple sentences and this allows to generate paragraph models. In contrast, an Objective (Hypothesis) can hardly generate a model, since typically has a sentence.

Our model incorporates four different schemes (methods) to assess conceptual sequences and each is applied depending on the section being evaluated. For *Problem Statement* and *Conclusions*, First Paragraph of Reference (FPR) and Evaluation of Nearest Neighbor (ENN) methods are appropriate. For the *Justification* section, Cascade Evaluation (CE), ENN, and Auto-evaluation of Paragraphs (AP) are pertinent.

The first step of the analyzer is preprocessing of the student writing, which is done in two tasks. The first focuses on segmenting each section into paragraphs. Each paragraph is defined as a sequence of sentences bound by the line feed. EGrid tool requires as input, the text in a Treebank format[1]. The second task is a process of translation from Spanish to English using Google Translator, since our corpora is in Spanish. The result of the translation enables processing the text with an English parser, in particular that of Stanford. Currently, the parsers for Spanish do not adhere to the Treebank tags, as English parsers. Since we are mainly interested in the entities, we just expect that the translator keeps consistent translating such entities, that in most cases occurred, given that they are mainly technical terms.

Our corpora for analysis and validation consisted of a total of 240 samples (theses), 40 samples for grad and 40 for undergrad level in each of the three sections. The first part included documents of master and doctoral level. The second part included documents of Bachelor and Advanced College-level Technician degree. We initially analyzed the behavior of transitions of the EGrid in 10% of samples of the graduate corpus, in the three sections, employing the tool developed by [5].

### 3.1 First Paragraph of Reference (FPR)

This method emerged after analyzing the *Conclusions* section. We observed that most of the transitions appear as first paragraph entities, i.e. entities identified by the EGrid tool in the first paragraph are further shown in the rest of paragraphs. Figure 2 illustrates this behavior. In addition, the remaining paragraphs also included same entities as the first paragraph. Note that the transitions appeared in a sequence (S, O, X, -).

[1] `http://www.cis.upenn.edu/~treebank/home.html`

For instance, in the marked rectangle, the entity "systems" was identified in the second sentence as subject. Afterward, in the third and fourth sentences, it appeared as the object, and finally was identified as object in the eleventh sentence. These transitions provide evidence that most paragraphs are adequately connected in term of concepts. The FPR scheme begins by generating a model of the first paragraph with the EGrid tool. Then, subsequent paragraphs are evaluated with such model, with the expectation that they get a positive value. The results provided by the EGrid tool are in the range from 0 to 1, where zero indicates a random (null) conceptual sequence. A result of one implies the existence of a relationship between the model and the evaluated new text, i.e. a subsequent paragraph, hence the conceptual sequences is strong.

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| THESIS | S | - | - | - | X | X | - | - | - | - | - | S | - | - | - |
| SOFTWARE | X | - | - | X | X | - | X | - | - | X | - | X | - | - | - |
| MODEL | X | - | O | O | X | X | - | - | X | - | - | - | - | - | O |
| SYSTEMS | X | S | O | O | - | X | - | - | - | - | O | X | - | - | - |
| TRANSFORMATION | X | - | - | - | X | O | - | - | X | - | - | - | - | - | X |
| TECHNIQUES | X | - | - | - | X | - | - | - | - | X | - | - | - | - | X |
| MODELING | - | X | - | - | - | - | - | - | - | - | - | - | - | - | - |
| STRATEGY-DO | - | O | - | - | - | - | - | - | - | - | - | - | - | - | - |
| DIAGNOSTIC | - | S | X | X | - | - | - | - | - | - | X | - | - | - | - |
| EXPERT | - | S | X | X | - | - | - | - | - | - | X | - | - | - | - |
| STRATEGY | - | - | S | - | - | - | - | - | - | - | - | - | - | - | - |
| MODELS | - | - | X | X | - | O | - | X | O | - | - | X | O | X | X |

**Fig. 2.** EGrid of a Conclusion (sample paragraph)

After comparing each result obtained by the EGrid of the graduate corpus with the content of each *Conclusions* section, we found adequate sequence in those paragraphs that showed a value higher than 0.5 and a weak sequence in those below that value. In assessing the subsequent paragraphs, we expect the results to be above zero, which would show that the *Conclusions* section have a suitable conceptual sequence.

### 3.2 Evaluation of Nearest Neighbors (ENN)

This method was designed to evaluate paragraphs that, after being examined with the FPR scheme, obtained a value of zero, i.e. did not show a conceptual connection. These were called "null paragraphs". The purpose of this method is to relate the null paragraphs with their close neighbor paragraphs, before and after. The first step is to evaluate the null paragraph with the prior neighbor paragraph model. If there is no relationship between them, we proceed to evaluate the paragraph with the subsequent neighbor paragraph model. Finding a relationship between the null paragraph and its neighbor paragraphs, we will detect a connection with the rest of the conclusions. Also, all paragraphs of the evaluated section would have some sequence of concepts. Otherwise, this paragraph has no relationship to the rest of paragraphs.

Figure 3a shows three paragraphs of a *Conclusions* section evaluated by the FPR method, with the first two having an acceptable sequence (strong connecting line).The third paragraph is null and shows no connection. When applying the ENN method on the null paragraph, we obtained a connection to its close prior neighbor (dotted line, Figure 3b), that enables the interconnection of the entire section and the null paragraph is actually related to the whole section.

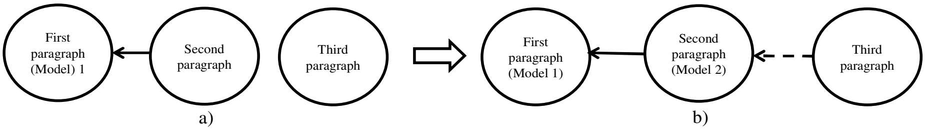



**Fig. 3.** a) Result after FPR, b) Result after ENN

### 3.3 Cascade Evaluation (CE)

This scheme evolved from what was observed in the EGrid regarding elements of the *Justification* section, i.e. it shows that transitions are distributed, and do not concentrate on one zone. A review between the EGrid and the evaluated original text allowed to find that some paragraphs presented a sequentially thematic relationship. For example, Figure 4 shows a partial EGrid of four paragraphs of a *Justification* of a graduate text. The third paragraph contained the entities "challenge" and "objective", identified as subject and object, respectively. Afterwards, the same entities appear as subjects in the fourth paragraph. This similarity in the two paragraph entities revealed that the paragraphs are indeed directly connected. This behavior looks like a thematic window between two paragraphs, showing a relationship to the preceding paragraph. This window moves across remaining paragraphs (Figure 4, see marked zone). The CE method works by generating the model of the first paragraph that is used to evaluate the second paragraph. Subsequently, we generate a model of the second paragraph, and the third is evaluated with this model, and so on for the rest of the section.

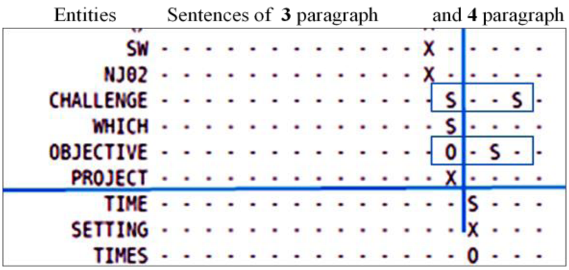



**Fig. 4.** EGrid of a Justification

### 3.4 Auto-evaluation of Paragraphs (AP)

This scheme was designed to evaluate null paragraphs that remained after being evaluated with any of the previously described methods. We decided to assess the paragraphs individually, hoping that they were at least connected by themselves, even though they were not related to the other paragraphs. The first step is to divide each of the paragraphs in two parts, but only those with at least four sentences (the EGrid tool does not generate models for one sentence). Afterwards, we apply the FPR method, i.e. generate the model of the first "paragraph" and then the second is evaluated with this model. In this way, we assess the connection within an individual paragraph.

Once we identified the main transition patterns, reaching the four schemes just described, we validate them in the remaining 90% of the corpus. This is reported in [6].

## 4 Conclusions

Assessing the sequence of concepts in proposal drafts is a complex task for computers, and often even for humans. We notice that the behavior of transitions in the *Conclusions* adhere to a pattern where most of the central entities concentrates in the beginning of the EGrid, i.e. the first paragraph contains information that was further developed in the other paragraphs. This behavior correspond to a pattern for conclusions[2], which begins with the restatement of the main premise, then a summarization of the key points, and finally the formulation of recommendations, assessments, and forecasts. The similarity between this pattern and that observed in the EGrid was verified by reviewing the texts of our corpus: we observed that when the student redefined the main problem, he used many of the key terms of his proposal, which are reflected in the subsequent paragraphs.

*Problem Statement* section showed a similar pattern as *Conclusions*. However, a slight difference in some cases was that most of entities appeared in the first two paragraphs. Regarding *Justification*, we notice that there exists a pattern where different entities are referred in a chained way, corresponding to different issues discussed at a time, as expected for this section.

The analyzer can be easily applied directly to English drafts, only by omitting the translation step, needed for Spanish. Moreover, drafts in other domains can also be analyzed without too much trouble. We expect that this analyzer will encourage students to develop better drafts and will contribute to an improvement in their writings.

**Acknowledgments.** Work supported by Conacyt (Schlp 1124002) and SNI, México.

## References

1. Davis, J., Liss, R.: Effective academic writing. The essay, vol. 3. Oxford University (2006)
2. Webber, B., Egg, M., Kordoni, V.: Discourse structure and language technology. Natural Language Engineering 18, 437–490 (2011)
3. O'Rourke, S., Calvo, R.: Analysing Semantic Flow in Academic Writing. In: Proceedings Conference on Artificial Intelligence in Education: Building Learning Systems that Care, pp. 173–180. IOS Press, Amsterdam (2009)
4. Barzilay, R., Lapata, M.: Modeling local coherence: An entity-based approach. Computational Linguistics 34, 1–34 (2008)
5. Elsner, M., Charniak, E.: Brown Coherence Toolkit. Brown University (2007), `http://cs.brown.edu/~melsner/manual.html`
6. Gonzalez-López, S., López-López, A.: Assisting Students in Writing by Examining How Their Ideas are Connected. In: Di Mascio, T., Gennari, R., Vitorini, P., Vicari, R., de la Prieta, F. (eds.) Methodologies and Intelligent Systems for Technology Enhanced Learning. AISC, vol. 292, pp. 9–18. Springer, Heidelberg (2014)

[2] `http://learninghub.une.edu.au/tlc/aso/aso-online/academic-writing/conclusion-paragraphs.php`

# Using JSON-LD and Hydra for Cloud-Based Tool Interoperability: A Prototype Based on a Vocabulary and Communication Process Handler for Mind Map Tools

Rocael Hernández[1], Christian Gütl[2], and Hector R. Amado-Salvatierra[1]

[1] GES Department, Galileo University, Guatemala
{roc,hr_amado}@galileo.edu
[2] IICM, Graz University of Technology, Graz, Austria
christian.guetl@iicm.tugraz.at

**Abstract.** Learning tools interoperability is a well-identified challenge tackled by different approaches such as the IMS Global LTIv2 definition. This paper presents a study to achieve interoperability in a cloud-based educational environment for external tools loosely coupled to form a virtual learning environment. This research builds on the W3C community group project Hydra to achieve interoperability for Web-service-based learning tools. A proof of concept is presented for one selected mind map tool supporting learning activities. The approach uses Web API descriptions based on a generic vocabulary built on Hydra and JSON-LD. A proposed communication process handler is presented in order to enable the interoperability. Initial results demonstrate that interoperability is highly efficient by means of a higher level of abstraction, while it simplifies the definition and coupling process of cloud-based tools.

**Keywords:** Cloud-based Tools, JSON-LD, Hydra, Interoperability, Cloud-based Educational Environment.

## 1 Introduction

Learning Tool Interoperability (LTI) is a highly relevant feature for further innovation in Virtual Learning Environments (VLE). It enables the use of new and specialized tools for the learning process in a single, unified, and seamless way. In this context, the IMS-Global LTIv2 specification [1] has become a widely recognized approach, one main feature is to VLEs to launch an external learning tool. It allows interaction with an external learning tool without further exchange of credentials, as well as interchange of context, roles, and other complementary educational data. LTIv2 [1] defines interoperability between the tool consumer (TC) (e.g., VLE) and the tool provider (TP) (e.g., Cloud-based Web 2.0 external tools). It supports the launch from the TC of a tool hosted by the TP, and it will exchange all necessary data to enable that tool to launch. The TP is also capable of storing data in the TC, associating that data with a course (i.e., a grade obtained in the TP can be transferred to the TC). This is reached by means of a definition of a standardized architecture for bi-directional

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 428–433, 2014.
© Springer International Publishing Switzerland 2014

communication that is based on RESTful [2] Web services. LTIv2 uses JSON-LD as its payload format. JSON-LD recently became a W3C standard. It is a serialization and messaging format, and through the use of linked data, interoperable Web services can be built. When implementing REST-based interoperability, a lack of use of Semantic Web features was identified in [3]. The general problem of scarce use of the semantic Web to design RESTful Web APIs was tackled by the Hydra W3C community group project [4][5]. The Hydra approach provides a means to describe Web APIs using a lightweight vocabulary and proposes the use of hypermedia controls to augment linked data. This initiative enables a TP to advertise valid state transitions to TCs in order to construct HTTP requests achieving truly RESTful services to improve interoperability. It also enables the TP to define the main entry point and to document the operations supported by the cloud-based tool (CBT).

This paper exemplarily presents a Web API description (in Hydra terms API Documentation or API Doc) for two selected mind map tools. The research is based on the Hydra generic vocabulary (GV) for mind maps presented in [6]. A Hydra GV describes the common resources, operations, and properties of a Web API, leaving tool-specific details to implementation time. This approach allows having an application domain type description of the Web API to avoid its duplication in each API Doc. The API Doc presented must be able to entirely describe the CBT Web API using the correspondent GV. To enable the interoperability with the aforementioned mind map tools, a communication process handler (CPH) is presented that acts as an intermediary layer between TC Hydra requests and TP standard non-Hydra responses. This solution will provide a lightweight infrastructure to process Hydra enabled requests and responses from CBTs (TP). Thus, the remainder of the paper is structured as follows: Section 2 gives an outline of a Hydra-based API Doc description for a CBT mind map, then in Section 3 a description of a proposed CPH for interoperability is presented. Section 4 presents a discussion of the findings, and finally the conclusions and future work are presented in Section 5.

## 2 API Doc Using Hydra GV for a CBT Mind Map

Various learning activities involve mind maps [7], a technique used to represent concepts' connections to other concepts. Mind map tools were exemplarily selected in this study to demonstrate the flexibility of a Hydra-based GV. From the existing mind map tools analyzed in [6], one of them was selected for this research: MindMeister. It provides a rich Web API covering mind map manipulation in the same way as using the MindMeister editor (i.e., create a map, add an idea, add relations between ideas, etc.). The GV presented in [6] contains the common properties, data types, and operations that a mind map tool may have. The three main resources are *user, map,* and *ideas*. Each resource has its own set of properties with its corresponding data types: mm:User (userId, userName, userEmail, maps), mm:Map (id, ideas, creationDate), and mm:Idea (id, title, imageUrl, modifiedBy, modifiedAt, parent). Each resource has its own set operations. All of them are subclasses of a hydra:Operation.

A Hydra API Doc describes all the current resources, operations, properties, and data types that a tool supports through its Web API. From the GV, an API Doc can use the common operations and properties related to a resource (a map, idea, or user). Thus, the API Doc is defined for each specific mind map CBT, and it uses the GV to define its own classes; for example, the API Doc defines a retrieveMap operation that is a subclass of the GV RetrieveMapOperation. The API Doc describes the IRI, HTTP method, and expected data and return values when the TC requests an operation. The API Doc from the MindMeister tool describes the three main resources of a mind map tool: *user*, *map*, and *idea*. A *user* resource (learners and professors in this context) uses the GV class User and the properties userId, userName, and userEmail. It describes operations and properties for these components. It also has one supported operation, which is retrievesUser, this gets all related information from a *user*. A *user* can be related to multiple *maps*, therefore the *user* property named User/maps is a collection of map resources that belongs to a *user*. That property also defines a set of supported operations related to it: retrieveMaps and createMap. Both operations use the GV RetrieveMapsOperation and NewMapOperation, respectively, but they define the correspondent HTTP method to be used and what it expects and returns in the request.

The *map* resource has the CRUD (Create, Read, Update, Delete) operations createMap, retrieveMap, editMap, and deleteMap, as well as some additional operations, such as shareMap and publishMap. Some operations use the GV. The *map* can have collaborators, which are a collection of *users*. This is supported by the operation retrieveCollaborators, which is a MindMeister-specific operation not described in the GV. The *idea* resource has the following properties: id, title, parent, modifiedAt, modifiedBy, imageUrl, and the operations addIdea and editIdea all use GV. There are also other operations and properties that are not part of the GV, such as correlations between ideas or positions of an idea in a graphical map.

## 3 Communication Process Handler for Web APIs and Hydra

As Hydra is a very new approach to JSON-LD-based Web service interoperability, it can be expected that currently available Web APIs do not make extensive use of it yet. However, many of the most mature and popular tools and services already have rich open APIs. That is the case for MindMeister, which has a RESTful XML-based API. Given that, a CPH has been implemented to invoke these services' APIs. Thus, communication can be based on Hydra, the mind map GV, and the specific API Docs created for each service, respectively. The current prototype permits a TC to invoke a mind map using Hydra APIs for MindMeister and Cacoo and processes their responses. The CPH request and response process model (RRM), consist of the following four steps: 1) a TC makes a request to the TP (the CBT); 2) the CPH has a collection of invoked service API method equivalents for each possible request, therefore it determines which is the correspondent method to perform the request for a given CBT; 3) the TP, such as MindMeister, offers responses in two formats: XML or JSON; and 4) the CPH with the payload obtained from the TP constructs the

Hydra-based response that the TC is able to process. It is worth mentioning that each Hydra response comes with the "@context" directive, which adds semantics to the response through referencing the API Doc and the corresponding data types to be used. This can initiate a discovery process of resources and their properties and operations, which can be applied recursively until information related to a given object (a map instance) has been obtained and processed. With a Hydra-enabled infrastructure capable of this discovery process, no specific coding for each tool is required. Only the Hydra GV and API Doc descriptions have to be used.

The CPH architecture maps the steps described in the CPH RRM: 1) TC makes Hydra request (Hydra API call receptor), 2) CPH sends request to TP (HTTP request maker and caller), 3) TP returns response (HTTP response processor and maker), 4) CPH returns Hydra response to TC. First a request receptor component handles requests to the CBT Hydra API, serving as proxy to the communication. Based on the TC request, the correspondent CBT API Doc is obtained. Then this is handled by the request processor, which extracts all related information that needs to go through for the final request to the CBT. This includes information such as expected payload, properties to be sent with the request, etc. The request will be formed based on the TC request but translated to the current Web API. That final request to the tool includes the use of Hydra *template links*, which serve to model automated URL queries to the TP. This gives the CPH a template to construct the URL at runtime, including the query parameters to be sent with the request. Once the HTTP request is made by the CPH, the *response processor* module takes the payload from the TP (CPH request and response model step 4) and validates the response. The correct format is verified for both JSON and XML responses. Then the response is Hydra formatted and serialized using the *response maker* module, and finally the CPH sends the response to the original TC. Some response properties from the CBT will not automatically match with the Hydra-defined properties that the TC requires, therefore a configuration interface module is provided to match all properties between API Doc and CBT. The property match process is done as an initial setup *contract* for a given TP, and it only needs to update that *contract* if the TP changes its response properties.

## 4 Discussion

The use case of mind map interoperability using JSON-LD, Hydra, and CPH has proven to be successful and a simple approach for education, in particular for VLE interoperability with CBTs. The VLE (TC) will need a Hydra interoperable infrastructure to be able to process requests and responses to the CBT. Such infrastructure is outside the scope of this paper, but it is worth mentioning that it will be in charge of semantically discovering and using the fine-grained operations available in the CBT. With that infrastructure, it will just be a matter of adding new CBT APIs, with no need to custom code interfaces for each new tool. The CBT that does not yet use Hydra will use the CPH as an intermediary to transform it to Hydra. The Hydra approach enables a high granularity management level of the VLE over the CBT, enhancing automation of the course administrative task over the tool (i.e., creating a map of each group in a class, assigning read permissions to comment on other groups' maps, etc.), and allows full access to the TP Web API. It only requires writing the Hydra definition of a Web API to include a new mind map tool. This can be achieved rapidly, without a large technical effort or complexity.

The CPH communicates back and forth with the cloud API by using Hydra definitions. Instead of writing code directly in the underlying API of a mind map (or any other tool, for that matter), all requests go through the CPH harmonizing the APIs. The advantage is that interoperability becomes much more efficient and the programming can happen on a higher level of abstraction. The whole system is built on a layered architecture that helps to concentrate the transformations on a single layer instead of having to spread them throughout the whole code base. As soon as the third-party APIs adopt JSON-LD and Hydra themselves, it becomes possible to completely eliminate the intermediary layer, which is the CPH. The Hydra approach is simpler because it just requires the CBT to publish its API using Hydra. It is not a specification of how interoperability should happen, as it is in the case of IMS LTIv2, therefore adoption and maintenance for the CBT providers is simpler. A Hydra interoperability infrastructure that can be plugged into the VLE will require no further coding to interoperate with each new CBT; in addition, it will enable fine-grained operations over CBTs that are not possible now with IMS LTIv2.

## 5 Conclusion and Future Work

Currently, to enable interoperability between a VLE and a Web API of CBTs requires custom-made interfaces for each tool to be integrated; that is, programming interfaces to each Web API. This implies a considerable effort to build and update each interface. The solution presented enhances that approach, as it converts Web APIs into APIs powered by JSON-LD and Hydra. It proves to be a robust way for the TC to control and administer tool resources, defining its properties and possible operations. At the same time, it simplifies interoperability efforts, enables a general abstraction for an application domain type (e.g., a mind map) through the GV, and permits the tool to be discoverable; for example, it finds all operations available for a tool at runtime without involving customized code.

The presented work simplifies interoperability for education and allows access to the complete set of features offered by a tool Web API, thus permitting it to create tailored learning experiences. It is relevant to mention that VLE will need an interoperability infrastructure to handle Hydra Web APIs. The API Doc presented in this work fully represents the CBT Web API, and the CPH correctly performs Hydra-enabled communication with the CBT Web API. The use of Hydra is a practical approach for interoperability, in contrast to a specification approach that requires TPs to adapt their APIs to specifications that are subject to change and might be a bit rigid [8]. With this interoperability approach, it is possible to enable learning orchestration mechanisms [9][10] that can enhance learner experience through the use of CBTs.

While TPs adopt Hydra, the CPH will play a major role, therefore a possible enhancement is to create an API Doc builder to simplify the developer's work to create and maintain the tool or service that he or she wants to convert to Hydra.

## References

1. IMS Global Learning Consortium, IMS GLC Learning Tools Interoperability Basic LTI Implementation Guide (2010)
2. Fielding, R.T.: Architectural styles and the design of network-based software architectures (Doctoral dissertation, University of California) (2000)
3. Lanthaler, M.: Model your application domain, not your JSON structures. In: Proc. of the 22nd Int. Conference on World Wide Web, pp. 1415–1420 (2013)
4. Lanthaler, M.: Hydra: A Vocabulary for Hypermedia-Driven Web APIs. In: Proceedings of the 6th Workshop on Linked Data on the Web (LDOW 2013) at the 22nd International World Wide Web Conference, WWW 2013 (2013)
5. Hydra W3C community group project, http://www.w3.org/community/hydra/ (last visit March 12, 2014)
6. Hernández, R., Gütl, C., Amado, H.: Interoperability for Cloud-based Applications for Education Settings based on JSON-LD and Hydra: Ontology and a General Vocabulary for Mind map Tools. In: Proc. of the 14th I-Know Conference. ACM (in press 2014)
7. Buzan, T.: The Mind map Book How to Use Radiant Thinking to Maximise Your Brain's Untapped Potential. Plume, New York (1993)
8. Henning, M.: The rise and fall of CORBA. Queue 4(5), 28–34 (2006)
9. Hernández, R., Gütl, C., Amado-Salvatierra, H.R.: Cloud Learning Activities Orchestration for MOOC Environments. In: Uden, L., Sinclair, J., Tao, Y.-H., Liberona, D. (eds.) LTEC 2014. CCIS, vol. 446, pp. 25–36. Springer, Heidelberg (2014)
10. Rizzardini, R.H., Linares, B.H., Mikroyannidis, A., Schmitz, H.C.: Cloud services, interoperability and analytics within a ROLE-enabled personal learning environment. Journal of Universal Computer Science 19(14), 2054–2074 (2013)

# The Design and Evaluation of a Sensor-Based Mobile Application for Citizen Inquiry Science Investigations

Christothea Herodotou, Eloy Villasclaras-Fernández, and Mike Sharples

Institute of Educational Technology, The Open University, UK
{Christothea.Herodotou,Eloy.Villasclaras-Fernandez,
Mike.Sharples}@open.ac.uk

**Abstract.** Despite their advantages of portability and ease of use, mobile devices have not yet been used in their full potential in education to measure and investigate real world phenomena. While some existing applications exploit individual sensors on mobile devices, there is no current toolkit that combines and customises data from the full range of sensors, and makes these data available for import to citizen inquiry science projects. This paper presents such a toolkit, called the Sense-it app, which gives access to all sensors on Android smartphones and tablets and connects to new or existing citizen science projects. We describe the design and formative evaluation of the toolkit in collaboration with students and teachers from a city technology college.

**Keywords:** mobile phone sensors, citizen science projects, mobile applications.

## 1 Introduction

The idea of having ordinary people act as citizen scientists has become increasingly popular during recent years. Non-professional members of the public can voluntarily collaborate with scientists in order to contribute data to natural science projects such as species identification and air pollution. Citizen science is an inexpensive way to collect large-scale data sets for research purposes and educate the public in scientific thinking. Yet, citizen science is facing challenges in extending to a young population disinterested in science education [1, 2]. Also, most citizen science projects take a top-down approach; apart from data collection activities, they do not engage volunteers in all the stages of the scientific process. The aim of this paper is to address these challenges through the design and evaluation of a mobile application (app), called Sense-it. Sense-it unlocks the hidden potential of smartphones by giving access to phone sensors and it links to a web-based platform (www.nquire.info) which hosts proposed investigations. The concept, design and example applications have been developed through a partnership with Sheffield University Technical College (UTC).

## 2 The State of the Art in Citizen Science

Citizen science is a term used to describe the research collaboration between scientists and the public in projects such as species recognition and counting, and water quality

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 434–439, 2014.
© Springer International Publishing Switzerland 2014

monitoring. The role of the public is to contribute to research-related tasks such as observation and measurement. The benefits of citizen science endeavours are scientific ("large spatial scales, long time series, data from private land, and labour-intensive data that would otherwise be expensive to collect") and educational including "educating the public in science and scientific thinking, inspiring appreciation of nature, and promoting support for conservation initiatives" [3, p.1]. Participation in citizen science projects could contribute to the demand for STEM proficiency by offering hands-on STEM opportunities to amateurs and boosting their interest in these disciplines [1].

One of the challenges citizen science faces is the perception that citizen science should be a means to recruit people as data collectors or analysts rather than engage them in all aspects of the scientific process. The frequently adopted model for doing citizen science is a top-down one with volunteers serving the traditional purpose of data collection [4]. Having non-professionals devise their own scientific questions and activities is a challenging task, for proposed investigations should be personally relevant, accomplished using recognised methods of data collection and analysis, valid and ethical [5].

## 3 Mobile Phone Sensing

Mobile phone sensing is a relatively new area of research related to the use of sensors embedded into smartphones, such as accelerometer and compass, and their application to domains such as social and environmental monitoring. Sensors can be organised into: a) motion sensors measuring acceleration and rotation using a 3-axis coordinate system (e.g., accelerometers), b) environmental sensors measuring environmental conditions such as air temperature (e.g., barometers), and c) position sensors measuring the physical position of the phone (e.g., GPS). Sensor-based applications designed explicitly for citizen science initiatives have been developed for the observation of nature such as *iSpot* (www.ispotnature.org), astronomy such as *Meteor Counter* (http://meteorcounter.com/), recording of environmental conditions (e.g., *NoiseTube*, http://noisetube.net/), and social science applications such as the *Community Ethnography app* (www.mtu.edu/news/stories/2013/october/story97909.html).

The main functionality of available mobile apps is to capture nature and wildlife and report on environmental conditions as a means to solve science-related problems. The role of the user is restricted to mainly collect and upload data to science websites where professional scientists take over and analyse data. Based on the type of citizen science project to be implemented, these applications make use of only a specific number and not the full range of sensors available on smartphones. To the best of authors' knowledge, no sensor-based mobile application has been yet developed which allows users to make use of all the sensors available on their mobile phones, connects sensor recordings to a diverse set of citizen science projects, provides instant feedback on how a sensor recording relates to other users' input, and scaffolds users in proposing and designing their own citizen science investigations. This paper presents the design of such an application, called the Sense-it app (available in the Google Play Store). It can be freely downloaded from Google Play under a 3-clause BSD licence.

# 4 Design of the Sense-it App

The Sense-it app is comprised of three tabs: the Explore, Record and Share tabs (see Figure 1). The Explore tab is used to preview a matrix of sensors available on a given smartphone. By touching a sensor icon, a graph is produced showing the live recording of the sensor. The Record tab is used to select sensors, manage their sampling rate, record and save data. Users have the option to preview a graph of the data during the period of recording, save and inspect the recorded data.


**Fig. 1.** The Explore, Record and Share tab

The use of Share tab is twofold. First, it allows users to create profiles using the sensors selected in the Record tab. Profiles are used to group and save a set of sensors for future use. For example, to map the noise in a neighbourhood, the sound and GPS sensors are selected. The sound sensor captures the level and sound density and GPS provides the location of the observation. A profile described as 'Noise map'can be created that includes the sound and GPS sensors. This profile is saved and is readily available when users access their application in the future. Second, it enables users to download existing profiles and contribute data to existing investigations by connecting to a web-based platform. Collected data can be uploaded to the web-based platform and automatically processed.

A web-based platform names nQuire-it has been created alongside the Sense-it app. Users can use the platform to propose and run citizen science investigations in informal settings, explore, visualize, and analyse their data and compare it to other users. Post-processing algorithms are applied to calibrate data from sensors, assist in data interpretation and integrate data from multiple recordings into a single plot. Instructions on how to author an investigation (e.g., how to record data using Sense-it) are given to ensure similar data acquisition conditions and improve reliability of collected data. The platform hosts a set of investigations including, 'Noise map' (Sound, GPS), 'How to attract bumble bees in your garden' (Camera, GPS) and 'Are birds scared by noise?' (Sound, Camera, GPS). Future work is underway to create a community of users online who will propose investigations that will make use of the full range of available sensors.

## 5 Evaluation of the Sense-it App

As a means to engage young people as informants [6] in the design process, we partnered with Sheffield UTC. The role of young people was to evaluate and improve the design of the tools, define the aims and design personally meaningful science investigations. The first prototype of the Sense-it app was evaluated by 96 students (16-18 years old) (Males = 86, Females = 10). Students were formed into 14 groups and they were asked to use and then evaluate the toolkit and propose science investigations that might be implemented using the toolkit. A set of worksheets was given to each group to support and guide their activities. Worksheet 1 asked groups to propose two science investigations by writing down a title, a specific question and how to use the app to collect data. Worksheet 2 asked groups to write down what they like the most about the app, what they like the least about it and what they would like to change on it.

Students' answers were analysed using thematic analysis. The answers were organized in clusters of relevant meaning and reduced into summary categories. The analysis revealed the following themes: (a) In terms of their evaluation of the application, students were found to be satisfied by the potential to customize sensors, the fact that the recording data was presented in two formats (numbers and graphs) and could be saved. Their suggestions for improvements were clustered around: accessibility issues including the complex information display on graphs and sensors, difficulty in navigation, the need for simplicity, and attractiveness in terms of colours, sensor icons design and fonts. (b) In terms of the proposed science investigations, these were grouped into the following categories: sound, light, acceleration and temperature. Some investigations suggested by the students are the following: How loud is it when you do Maths compared to English?, How bright does light need to be to wake you up?, What is the acceleration and top speed of the lifts in UK?

The review of students' proposed investigations revealed that the identification of science investigations that can be easily implemented and of interest to the majority of young people is a rather difficult task [7]. The investigation 'Find the fastest lift' proposed by the students has already been integrated into the Sense-it app. Based on students' recommendations, the following changes were made to the app: depictive icons for each sensor were designed, a help button was added giving a brief description of how each sensor works, and the three tabs of the app were renamed into Explore, Record, Share pointing to their main functionality.

## 6 Discussion

The development and evaluation of a sensor-based application which combines and customises data from the full range of phone sensors and makes these data available for import to citizen science projects comprised an important step towards engaging young people in doing citizen science. Initial insights from the evaluation of the Sense-it app reveal that young people liked the fact that they can use their phones and tablets to do science as well as the design of a flexible app which unlocks their phones potential.

The active participation of young people to the process of development, evaluation and use of the Sense-it app is an example of how citizen science projects could shift from the frequently adopted top-down model of volunteers serving only purposes of data collection [4]. Young people had the opportunity to negotiate how the application looks like and functions to better fit their expectations. In addition, the flexibility of the application offered young people the opportunityto set up and run their own science investigations and get credit for their contributions by assigning the authoring of each investigation to the person/s who proposed it in the web-based platform. As a final step, new investigations will be revised by experts to ensure young people provide detailed instructions on how data would be collected thus enhancing their quality.

## 7 Conclusions and Future Directions

We have described a mobile application which uses all the sensors on mobile devices, links to diverse citizen science projects, provides instant feedback on how sensor recordings relate to other users data, and allows users to create their own science investigations. The next step is to analyse emerging learning gains related to the scientific method and evaluate users' understanding of instructions on how to use Sense-it to run investigations and visualize and compare data.

Although initial findings from use with young people revealed that they are interested in using the Sense-it app and defining their own personally relevant science investigations, there still remains an issue of how to sustain their interest and motivate long-term engagement and participation in citizen science activities. Among the mechanisms to be considered in future studies is web 2.0 technologies, for increasing popularity can result in increasing adoption [2], game-based mechanisms such as missions and rewards, and online community creation.

**Acknowledgements.** This work is funded by Nominet Trust. The authors would like to thank students and staff from Sheffield UTC and members of The Open University for participating in Sense-it evaluation.

## References

1. Toerpe, K.: The Rise of Citizen Science. The Futurist, 25–40 (July-August 2013)
2. Catlin-Groves, C.L.: The Citizen Science Landscape: From Volunteers to Citizen Sensors and Beyond. International Journal of Zoology 2012, Article ID 349630, 1–14 (2012)
3. Freitag, A., Pfeffer, M.J.: Process, not product: investigating recommendations for improving citizen science "success". PloS One 8(5), 1–5 (2013)
4. Mueller, M., Tippins, D., Bryan, L.: The Future of Citizen Science. Democracy and Education, 20, 1 -11 (2012)
5. Villasclaras-Fernandez, E.D., Sharples, M., Kelley, S., Scanlon, E.: Supporting Citizen Inquiry: An Investigation of Moon Rock. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 383–395. Springer, Heidelberg (2013)

6. Scaife, M., Rogers, Y., Aldrich, F., Davies, M.: Designing for or designing with? Informant design for interactive learning environments. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 1997), pp. 343–350. ACM, New York (1997)
7. Anastopoulou, A., Sharples, M., Ainsworth, S., Crook, C., O'Malley, C., Wright, M.: Creating personal meaning through technology-supported science learning across for-mal and informal settings. International Journal of Science Education 34(2), 251–273 (2012)

# DDART, a Dynamic Dashboard for Collection, Analysis and Visualization of Activity and Reporting Traces

Min Ji[1], Christine Michel[1], Élise Lavoué[2], and Sébastien George[3]

[1] Université de Lyon, CNRS
INSA-Lyon, LIRIS, UMR5205, F-69621, France
{min.ji,christine.michel}@liris.cnrs.fr
[2] Magellan, IAE Lyon, Université Jean Moulin Lyon 3, France
elise.lavoue@univ-lyon3.fr
[3] LUNAM Université, Université du Maine, EA 4023, LIUM, 72085 Le Mans, France
Sebastien.George@univ-lemans.fr

**Abstract.** Most Learning Management Systems (LMS) focus on providing rich learning materials to the learners but rarely offer possibilities to monitor and analyze their learning processes. However metacognitive strategies argue that learners can improve their learning performances by monitoring their activities, especially in the context of Project-Based Learning (PBL). Our work consists in designing tools to support metacognitive strategies. In this paper, a framework and a dynamic dashboard for PBL situations are proposed to help learners to collect, analyze and visualize the meaningful traces of their activities by themselves. The proposed framework integrates activity traces (interactions with the LMS) and reporting traces (semi-structured sentences written by the learners). The dynamic dashboard supports learners in the creation customizable indicators through a user-friendly interface.

**Keywords:** project-based learning, indicator, dashboard, self-regulation, metacognition.

## 1 Introduction

The World Wide Web provides opportunities for creating virtual classrooms for learners and instructors [1]. Unfortunately, most Learning Management Systems (LMS) rarely offer possibilities to monitor and analyze their learning processes to improve their learning performances. Hence, our research aims to improve learners' meta-cognition and self-regulation awareness by exploring the traces produced during the learning processes. We focus on the development of tools to help learners to monitor their own learning activities.

Our works are conducted in the context of Project-based-Learning (PBL). Michel *et al.* [2] found that the use of a dashboard in PBL could help the learners to develop metacognitive skills and new behaviors. Based on this fact, our general research question is how to design a dashboard to help learners to self-regulate and build metacognitive skills during their project activities.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 440–445, 2014.
© Springer International Publishing Switzerland 2014

In our previous works[3, 4], we have defined a dashboard used in the PBL context as a Project Based Learning Management System (PBLMS). We have proposed a general architecture based on the management of the traces of the project activities. This architecture combines activity traces and reporting traces into a dashboard to support the learning processes [4]. The activity traces can reflect the interactions in the PBLMS environment, while the reporting traces can record the interactions out of the PBLMS environment. These two types of traces can give a general view of the project and can support self-regulation processes specifically if the learners are able to explore them directly.

In this paper, we propose an operational dynamic dashboard for learners using both activity and reporting traces. In part 2, we introduce the existing dashboards used to support self-regulation and give a critical analysis. In part 3 we present the DDART system, a dynamic dashboard we developed. We finally conclude in part 4 and detail our future works.

## 2 State of the Art

A self-regulating dashboard should present the information about goals and activities at a glance on the shape of indicators that allow easy navigation to more complete information on analysis views [5]. So, the dashboard could be considered as a container of indicators [3]. In order to calculate the indicators, we have identified two types of traces useful to build the indicators: activity traces and reporting traces [4]. This part describes the existing simple dashboards that enable the users to use predefined indicators (all the aspects of the indicator are defined by the system developers totally) and dynamic dashboards that let the learners customize their own indicators (users can define all the aspects of indicators).

### 2.1 Simple Dashboards with Predefined Indicators

Simple dashboards are designed to help the users to monitor some specific aspects of their activities. CourseVis [1] computes graphical indicators by activity traces in WebCT. TrAVis [6] is a reflective tool that also uses activity traces to give students information about the way they carry out discussions or other collaborative activities. Students have access to the tracking data repository. Study desk [7] supports the monitoring of the learners' learning processes by evaluating the progresses and tasks from activity traces. Feeler [8] uses personal physical data to reflect the relationship between learners' learning performances and well beings with the aim of fostering reflection and awareness. NAVI Badgeboard [9] aims at improving individual awareness and reflection of personal activities through visualizing learners' communication activities by "badge" presentation.

The system gStudy [10] is a collaborative learning platform. The traces collected are the log files and the messages in the module gChat. All the traces are analyzed by the module LogAnalyser. Radar [11] is a peer feedback tool that provides users with anonymous information on how their cognitive and social behavior is perceived by themselves, their peers and the group as a whole. Reflector [11] is a reflection tool that computes learner's answers about five reflective questions. PCO-vision [2] is a dashboard that supports self-regulation by offering a global view on objectives-actions-results. The learners self-declare their goals and the way they carry out activities. They can explore these reporting traces to self-regulate themselves.

### 2.2 Dynamic Dashboards with Customizable Indicators

Dynamic dashboards offer three customization levels for the indicators. In a low level, users can set some simple parameters of the indicators; in a medium level, users can define the calculation functions and change the visualizations of the indicators; in a high level, learners can manipulate the traces used in the indicators. Navi Surface [9] has a low level of customization. It uses the same principle of badges presented with NAVI Badgeboard (mentioned in last section). Users can specify the names of the group members in order to view how the group badges acquisitions arise. The only visualization mode is badges. The Academic Analytics Tool (AAT) [12] supplies tutors with medium customization functions by allowing them to extract specific information from the activity traces and to select the calculation methods through a SQL query GUI. Specific computer skills are needed (write SQL) and the only visualization mode of the results is tables. GINDIC [13] supports the generation of high customizable indicators by using activity traces but requires the users to have a computer background because many parameters need to be set. The TBS-IM system [14] uses the concept of modeled trace (M-Trace) to enable tutors to select, transform and visualize activity traces produced by Moodle. The customization process is high. It offers various visualization modes to users but tutors are required to have some computer skills to understand how to manipulate the traces and to define parameters.

### 2.3 Summary / Critical Analysis of Existing Dashboard

The previously presented dashboards can be compared from four aspects: target user, trace type, customization level and computer skills requirements.

- Most dynamic dashboards are dedicated to tutors and researchers while most simple dashboards are supplied to learners. Learners have no means to self-regulate their activities and to build metacognitive skills.
- Most dashboards explore the activity traces to analyze the learning processes while the reporting traces are ignored. The whole learning processes can be observed completely only by combining the two types of traces.
- Most dashboards don't supply the customization functions or just supply some low customization functions.
- The higher the customization level is, the higher the computer background requirements are. There is no simple but high-dynamic dashboard for learners.

Based on the above analysis, we propose in the next section a dynamic dashboard based on the LMS Moodle.

# 3 DDART System

DDART is a Dynamic Dashboard Based on Activity and Reporting Traces. This system allows learners to create customizable indicators related to their activities. We choose to implement it as a plug-in of the Moodle platform. It is based on a four steps framework detailed in each part of this section: data collection, data integration, data calculation and data visualization.

## 3.1 Data Collection

Two types of traces are collected: activity traces and reporting traces. We have classified the activities in PBL into three categories: internal activities, external activities and non-instrumented activities [4]. We extract parts of the internal activities data, recorded directly by Moodle, to constitute the activity traces. These traces come from the use of five modules integrated in the platform (wiki, chat room, forum, private message and resource) and of two modules we developed: a reporting tool and DDART. The reporting traces are recorded when learners write reports in the reporting tool [4] to declare their non-instrumented activities and external activities. The reports are composed of semi-structured sentences.

## 3.2 Data Integration

This step aggregates the two types of traces together and generates a primary trace (PT). A primary trace is defined as "*a trace model and a set of untransformed traces instances according to this model*" [14]. It is the source data for creating the indicators. Based on our PBL context, we propose a task-oriented pattern model to store the primary traces:

$$\mathrm{PT} = \{\mathrm{Ta}, \mathrm{L}, \mathrm{C}, \mathrm{V}, \mathrm{To}, \mathrm{BT}, \mathrm{ET}, \mathrm{P}\} \tag{1}$$

With:

- Ta: the task carried out by learners during the project; L: the learner who carries out the task; C: the category of the task; V: the value (detailed content) of the task; To: the tools supplied in Moodle to help learners to complete the project; BT and ET: begin time and end time of the task; P: the place at where the task was done.

A filter function is supplied to help learners to set the constraints of learners, time, places, tools and tasks so as to have a smaller set of traces.

## 3.3 Data Calculation

The Fig. 1 presents the interface of the indicator design process. This interface is composed of three parts: (1) the "parameters" part ( see Fig. 1.a ) contains the list of all the parameters which are available for creating an indicator, (2) the "calculation" part ( see Fig. 1.b and c ) allows learners to specify the parameters and presents the view of the indicator results and (3) the "visualization modes" part ( see Fig. 1.d ) supplies ten presentation models to learners to choose.

This user-friendly interface is WYSIWYG (what you see is what you get) and the results can be calculated on the fly, without delay, so that learners can adjust the parameters all the time. It allows learners to create the indicators by dragging and dropping the parameters and the visualization mode. Learners can manipulate five entities E={L, To, Ta, Time, P} from the PT model. They can cross the entities together, select the type of value to consider (frequency, time interval, content and description) or do some mathematical calculations (sum, means, etc.). The cross operator calculates the cartesian product E1×E2 as follow:

$$E_1 \times E_2 = \{E_{11}, \cdots, E_{1n}\} \times \{E_{21}, \cdots, E_{2m}\} = \{E_{11}E_{21}, \cdots, E_{11}E_{2m} \cdots, E_{1n}E_{21}, \cdots, E_{1n}E_{2m}\} \quad (2)$$

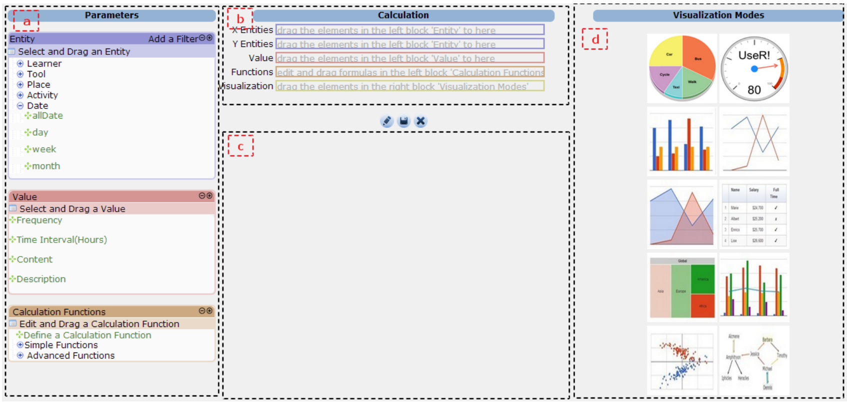



**Fig. 1.** The main page for trace transformation and indicator calculation

### 3.4 Data Visualization

DDART offers ten visualization modes for learners: pie chart, bar chart, line chart, gauge chart, social network, scatter chart, area chart, table, tree map and combo chart. We believe that different visualization modes can help learners to obtain different information from different dimensions. We imported Google Chart API to generate these visualizations automatically. All the created indicators are presented in the learners' own dashboards. Learners can follow their project and learning by observing their own customizable dashboards.

## 4 Conclusion

In this paper, we have presented the DDART system, a dynamic dashboard integrated to Moodle. This system can help learners to collect, analyze and visualize their traces in the form of meaningful indicators. By providing learners with DDART, we encourage them to explore information by themselves and learn how to regulate their learning activities. Considering our future research work, we will conduct an experiment in a real PBL situation to test the usability and the utility of DDART as a support for learners' self-regulation.

## References

1. Mazza, R., Dimitrova, V.: CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. International Journal of Human-Computer Studies 65(2), 125–139 (2007)
2. 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)
3. Ji, M.: Ergonomic Study of Existing Project-Based Learning Management System. In: The Fourth Young Researchers Conference on Information Environment for Human Learning, pp. 57–62 (2012)
4. Ji, M., Michel, C., Lavoué, E., George, S.: An Architecture to Combine Activity Traces and Reporting Traces to Support Self-Regulation Processes. In: IEEE 13th International Conference on Advanced Learning Technologies (ICALT), pp. 87–91. IEEE (2013)
5. Michel, C., Lavoué, E.: KM and Web 2.0 Methods for Project-Based Learning. MEShaT: a Monitoring and Experience Sharing Tool. Multiple Perspectives on Problem Solving and Learning in the Digital Age. p. 49-66. Ifenthaler D., Isaias P., Spector J.M., Kinshuk, Sampson D., Heidelberg (2011).
6. May, M., George, S., Prévôt, P.: TrAVis to Enhance Students' Self-monitoring in Online Learning Supported by Computer-Mediated Communication Tools. Computer Information Systems and Industrial Management Applications 3, 623–634 (2011)
7. Narciss, S., Proske, A., Koerndle, H.: Promoting self-regulated learning in web-based learning environments. Computers in Human Behavior 23(3), 1126–1144 (2007)
8. Durall, E., Toikkanen, T.: Feeler: Feel good and learn better. In: Third International Workshop on Awareness and Reflection in Technology-Enhanced Learning (ARTEL 2013), pp. 83–89 (2013)
9. Charleer, S., Klerkx, J., Santos, J.L., Duval, E.: Improving awareness and reflection through collaborative, interactive visualizations of badges. In: Third International Workshop on Awareness and Reflection in Technology-Enhanced Learning (ARTEL 2013), pp. 83–89 (2013)
10. Winne, P.H., Hadwin, A.F., Gress, C.: The learning kit project: Software tools for supporting and researching regulation of collaborative learning. Computers in Human Behavior 26(5), 787–793 (2010)
11. Phielix, C., Prins, F.J., Kirschner, P.A.: Group Awareness of Social and Cognitive behavior in a CSCL Environment. In: 9th International Conference of the Learning Sciences. International Society of the Learning Sciences. International Society of the Learning Sciences, vol. 1, pp. 230–237 (2010)
12. Graf, S., Ives, C., Rahman, N., Ferri, A.: AAT: A tool for accessing and analysing students' behaviour data in learning systems. In: The 1st International Conference on Learning Analytics and Knowledge, pp. 174–179 (2011)
13. Gendron, É.: Cadre conceptuel pour l'élaboration d'indicateurs de collaboration àpartir des traces d'activité. University of Claude Bernard Lyon I (2010)
14. 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)

# Labeling Mathematical Errors to Reveal Cognitive States

Thomas S. McTavish and Johann Ari Larusson

Center for Digital Data, Analytics, and Adaptive Learning Pearson, Austin, TX, USA
{tom.mctavish,johann.larusson}@pearson.com

**Abstract.** While technology can enhance learning, ironically, many online systems occlude learners' cognitive states because instructors do not directly observe students solving problems. In this paper, we show how we utilized an online mathematics homework system where students simply provided final answers to exercises. We then asked, "What can we infer about the cognitive state of the student if they gave an incorrect response?" Through data mining techniques, we found we were able to ascribe a particular type of mechanical error or misconception to 60-75% of the incorrect responses learners made on the subset of problems we analyzed. As such, we illustrate methods for extracting this data to discover knowledge components embedded in an exercise, expose item bias, and reveal learners' cognitive states.

## 1 Introduction

In education, "going digital" runs a gamut. The low end consists of digital forms useful for analytics, search, and archival. The high end consists of complete monitoring of learner interactions along with their biometric measurements [6]. While the high end of the spectrum offers the finest resolution of student learning, it also requires the learner to be, in a sense, "jacked in" and constantly surveilled, which may impede learning [9].

In this paper, we targeted the lower end of the digital spectrum by working with an online mathematics homework system where students simply entered a final, free text response to each exercise. We then asked, "What can we infer about the cognitive state of the student when they answered incorrectly?" Even without intermediate steps embedded in the task we were still able to label 60-75% of the errors in various exercises. We used relatively large sample sizes and exploited templates, each instance of the template having different parameters (e.g., "4+7" or "3+5"), which permitted us to see repeated patterns. We ascribed several types of errors to even seemingly straightforward exercises, demonstrating that several Knowledge Components (KCs) – the skills, concepts, and rules embedded in an activity [8] – may go into an exercise. Additionally, we were able to determine bias in the template that made the problem easier or more difficult depending on the parameters given. Methods such as ours illustrate approaches to inferring learners' cognitive states without complex assessment design and without perpetual interaction and surveillance.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 446–451, 2014.
© Springer International Publishing Switzerland 2014

## 2 Methods

Our data was comprised of user responses to online quizzes and tests using math exercises from a college level developmental math book. Responses were free text, final answers and graded as "correct" or "incorrect". The system randomly creates instances of each template using automatic item generation, often with certain constraints in the hopes of keeping the problem within the same domain and similar range of difficulty. In the two examples in this paper, one had 51 instances and the other had 1022.

We gathered student responses from each exercise. Since the response field is free text, it was straightforward to find those cases where several students converged to enter the same incorrect response. To determine *how* students arrived at their particular response, we looked across the instances of the template, comparing the peak responses to determine consistent patterns. After determining the type of error, we wrote mathematical expressions to match the specific error for the given input parameters of the instance. We therefore tagged those responses that had one or more errors attributed to them. (Since an incorrect response might match more than one error formula, the "Total inferred" row at the bottom of many of the tables we provide in our results is not a subtotal of each error category. Each tagged incorrect response was only counted once.) We then filtered out these cases and iteratively considered the remaining responses in attempts to further ascribe possible types of errors.

We determined bias in a template in the following three ways: We first considered the fraction correct with a particular instance as compared with the rest of the instances in a binomial test. We then looked at all instances that had a variable set to a specific value. Taking each variable across all of its values and performing the same binomial test allowed us to determine if and when a variable showed bias. We carried this one step further and performed the same binomial test on pairs of instance variables as well.

## 3 Results

### 3.1 "Find a Quotient" Example

Students were presented with the exercise "What is the quotient of $x$ and 5?" where $x$ was a uniformly random multiple of 5 in the range $[1000, 1250]$. We evaluated 2467 responses of which 379 were incorrect. With 51 possible instances with $x \in \{1000, 1005, \ldots, 1245, 1250\}$, there were only 7 incorrect responses per instance on average. Nevertheless, we noted that 31% of the errors followed the form $x \times 5$ and 12% of errors matched $x + 5$, indicating a misconception or misunderstanding surrounding "quotient". Interestingly, while many errors either multiplied or added, less than 1% of students gave a response that matched $x - 5$, implying that these students realized "quotient" did not involve subtraction.

We evaluated a similar exercise, simply stated as:

$$\text{Divide:} \quad x\overline{)y}$$

(Type the whole number part and the remainder part of the quotient. Type 0 in the second answer box if there is no remainder.)

Users were provided with two text fields to type the whole number and remainder. This format and set of instructions with different conditions for $x$ and $y$ comprised many exercises of the preceding section in the book relative to the "Find a quotient" example. In this particular exercise $x$ was an integer in the range $[2, 9]$ and $y$ was an integer in the range $[1000, 5000]$, calculated such that there was no remainder, so users were always expected to give "0" as the remainder. The errors from this long division problem never followed the format of those errors just described in the "Find a quotient" example. For example, students never expressed the answer as $x \times y$. In fact, of those students who solved both exercises, 57% of those that got the "Find a quotient" example wrong had previously solved the long division problem correctly. Collectively, it implies that many students simply stumble on the definition of "quotient" or were thrown by the change in format to the question.

With the remaining responses to the "Find a quotient" example, we could see that if the correct answer contained a "0" digit, that many student responses omitted it. For example, if users were asked to find the quotient of 1015 and 5, which is 203, they might give "23". Such a response indicates a mechanical error or misconception surrounding place values and accounted for 10% of all errors. In fact, while the mean probability correct was 85% for this problem, when contrasting problems that had a "0" in their correct answer vs. those that did not, the probability of success was 81% ($p = 0.008$) and 87% ($p = 0.025$), respectively, indicating that the problem added another knowledge component when asking students to deliberately consider the place value. Collectively, we could characterize 58% of the errors from this problem, which are summarized in Table 1.

**Table 1.** Summary of errors to the "Find a quotient" example

| ERROR | FREQ (%) |
|---|---|
| numerator × denominator | 31 |
| numerator + denominator | 12 |
| Omitted "0" in answer | 10 |
| Simply gave denominator | 3 |
| Simply gave numerator | 2 |
| numerator − denominator | 1 |
| Total inferred | 58 |

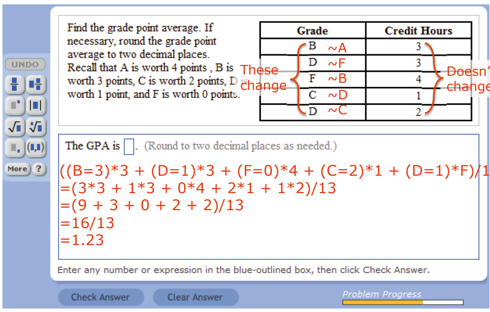



**Fig. 1.** Exercise to find the grade point average (GPA). Each row of grades in the left column can be an A, B, C, D, or F, except as specified. The credit hours column is fixed.

### 3.2 "Calculate a GPA" Example

As another example, we used the problem shown in Figure 1, which asks students to calculate a GPA. A nice feature of this template is that the *Credit Hours* column stays fixed. As such, all correct answers divide the sum of the grade points by 13. With 6322 exposures and a 43% success rate, we also had a large sample of errors ($n = 3634$) to work with. Since one grade letter was not permitted for each row, there were 46 possible instances such that correct answers will divide some numerator $\{3, 4, ..., 48, 49\}$ by 13 and round to the nearest hundredth. We contrasted the correct and incorrect answers. Interestingly, several incorrect responses were given as a number properly divided by 13 and rounded. In this case, learners obviously calculated an incorrect numerator. This accounted for 24% of the errors. Another interesting pattern we observed was that many incorrect responses ended in .0, .2, .4, .6, or .8. In other words, learners were dividing by 5, the number of classes, instead of 13, the number of credits. This misconception revealed itself in a few ways as enumerated in Table 2.

We also observed a lot of bias in this template. Unsurprisingly, if the GPA to be calculated was 2.0 or 3.0, the problem was significantly easier (50% fraction correct) than if it was $12/13 = 0.92$ (27% correct, and the most significant response). We also observed that if three "Fs" were presented, the fraction correct dropped to 31% ($p = 6 \times 10^{-5}$). While four "Fs" had a success rate of 27%, it was not statistically significant because of the few samples. We also found that when "As" were absent, the fraction correct was 39% ($p = 0.0016$) and when two "As" were given the success rate went up to 46% ($p = 0.016$). Collectively, these results hint that students understand the value of an "A" more than other letters, especially "F". Again, we speculate that perhaps students found it difficult to imagine taking a course and not receiving any credit for it.

## 4 Discussion

In this paper, we showed how incorrect, final responses to mathematics exercises can be aggregated and analyzed to infer specific types of errors. We were able to label 60-75% of the errors in the exercises we sampled and ascribe at least a half dozen error types to each problem. This work simultaneously quantifies intra-template variability, possible knowledge components in an exercise, and enables students' cognitive states to be inferred by labeling their errors.

### 4.1 Shortcomings and Promise of Technology

The most time-consuming task, while also being the most critical component of teaching, is the instructors' continual assessment of student performance [4]. Students are doing more and more online, so instructors do not directly observe students solving problems. In many cases, as in systems such as the one we used where students simply enter a final answer, instructors do not see worked solutions. Furthermore, as convenient as automatic scoring may be, many scoring

**Table 2.** Summary of errors for the GPA problem. "Summed grades" are where students applied $A = 4$, $B = 3$, $C = 2$, $D = 1$, and $F = 0$ to sum their set of grades, but did not multiply by the credit hours.

| ERROR | FREQ (%) |
|---|---|
| Wrong numerator (Counting error) | 24 |
| Summed grades, but did not multiply by credit hours, then divided by number of classes instead of credit hours | 21 |
| Answer out of bounds > 4.0 | 13 |
| Rounding errors | 11 |
| Summed credit hours, but did not multiply by any grades, then divided by the number of classes (13/5 = 2.6) | 10 |
| Summed grades and divided by the grade points | 5 |
| Correct numerator, but divided by number of classes instead of number of credit hours | 5 |
| Summed grades, but did not multiply by credit hours (subset of "Wrong numerator") | 2 |
| Did not credit "Fs" | 1 |
| Correct numerator, but never divided | 1 |
| Summed grades and never divided | 1 |
| Total inferred | 75 |

systems do not incorporate subtleties. More often than not, an item is deemed correct or incorrect and scored by the item's weight.

Our approach offers a window into students' cognition that is often occluded by technology. We show ways of discriminating misconceptions from mechanical errors, which can then be communicated to instructors and students. The system itself can score items differently and adapt to the learner. The most critical component of successful learning is prompt, detailed, positive and timely feedback on student work [2]. Indeed, while technology may remove certain face-time with instructors and a given system my provide insufficient metrics and feedback to instructors and students, technology also has the capability of providing an informative, data-rich environment [3].

### 4.2 Bug Libraries Revisited

Analytics surrounding mathematical errors is hardly new. Our work is reminiscent in spirit to the "bug library" work performed by VanLehn and Brown [1,7]. They constructed libraries of incorrect repairs to impasses to simple mathematics problems. Their approach combined cognitive models to algorithmically generate possible bugs and extensive analysis on handwritten mathematics problems solved by students. They ascribed, for example, over 70 bugs for subtraction. As the complexity of mathematics problems grows, constraining the error space to a limited, computational domain is nearly impossible. Automated methods to infer bugs are still needed. Nevertheless, we demonstrate some analytic methods and considerations that might be utilized in the construction of automated methods directed at complex problems. For example, the "calculate a GPA" exercise had a few important design characteristics. For one, it kept the *Credit Hours* column fixed. This made a discrete set of correct solutions by which the incorrect solutions could be contrasted. Furthermore, two of the variables – the number of classes and the total credits – were each a prime number. Therefore, the incorrect answers were, in a sense, more easily parsed. If the number of credits were 10, for example, then it might not have been as apparent that many students were dividing by the number of classes instead of the number of credits.

### 4.3 Knowledge Components and Bias

Knowledge components may be explicitly incorporated by task authors, or statistically revealed if unknown [5]. By labeling errors, we illuminate the KCs that have failed, demonstrating another means of discovering them. Template bias also indicates the KCs at play. We saw in our "Find a quotient" example that students often omitted a "0" in their answer if the correct solution included it. Such bias revealed a "0 in the tens place" KC that was absent in the other instances. Likewise, our GPA example showed more students struggling with "Fs", incorrectly rounding, or having problems with the division when the GPA was not a whole number. This revealed that different KCs were being challenged with different instances. Collectively, our work demonstrates means of capitalizing on digital approaches even when much work may be performed offline.

## References


1. Brown, J.S., VanLehn, K.: Repair theory: A generative theory of bugs in procedural skills. Cognitive Science 4(4), 379–426 (1980), http://onlinelibrary.wiley.com/doi/10.1207/s15516709cog0404_3/abstract
2. Brown, S., Race, P., Rust, C.: Using and experiencing assessment. Assessment for Learning in Higher Education, 75–85 (1995)
3. Conole, G., Warburton, B.: A review of computer-assisted assessment. ALT-J 13(1), 17–31 (2005), http://journals.co-action.net/index.php/rlt/article/view/10970
4. Crooks, T.J.: The impact of classroom evaluation practices on students. Review of Educational Research 58(4), 438–481 (1988), http://rer.sagepub.com/content/58/4/438, doi:10.3102/00346543058004438
5. Junker, B.W., Sijtsma, K.: Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. Applied Psychological Measurement 25(3), 258–272 (2001), http://apm.sagepub.com/content/25/3/258, doi:10.1177/01466210122032064
6. Picard, R.W.: Affective Computing. MIT Press (2000)
7. VanLehn, K.: Mind Bugs: The Origins of Procedural Misconceptions. MIT Press (1990)
8. VanLehn, K.: The behavior of tutoring systems. International Journal of Artificial Intelligence in Education 16(3), 227–265 (2006), http://iospress.metapress.com/content/AL6R85MM7C6QF7DR
9. Wang, F., Hannafin, M.J.: Design-based research and technology-enhanced learning environments. Educational Technology Research and Development 53(4), 5–23 (2005), http://link.springer.com/article/10.1007/BF02504682

# Responsive Web and Adaptive Web for Open and Ubiquitous Learning

Marco Mercurio[1], Ilaria Torre[1], and Simone Torsani[2]

[1] Department of Computer Science, Bioengineering, Robotics, University of Genoa, Italy
[2] Department of Modern Language and Culture, University of Genoa, Italy
{ilaria.torre,marco.mercurio,simone.torsani}@unige.it

**Abstract.** Online educational processes make use of different tools and platforms through which learning resources can be shared. Nowadays mobile platforms are gaining momentum and learners can experience mobile ubiquitous learning. This heterogeneity of resources and devices makes the design of e-learning content a relevant issue. In this direction, approaches to design websites and applications that are responsive and adaptive to different kinds of devices in different contexts is growing in importance. In this paper we analyze this issue and discuss both the responsive and the adaptive approaches to e-learning content design in the domain of language learning. Through an experimental evaluation, we will try to demonstrate that, given the nature of language learning, responsive web design approach alone seems not sufficient to provide an effective adaptation of content, especially of learning activities, to fit the device features, while adaptive techniques have better chances to.

**Keywords:** language learning, adaptive systems, responsive web design, human-computer interaction, mobile learning, ubiquitous learning, semantic web.

## 1 Introduction and Background

With the spread of mobile technology, education is evolving toward a model in which learning is a pervasive and continuous process, more and more embedded within everyday digital activities. An obstacle to open and ubiquitous learning, however, is that mobile devices may have several different hardware and software features. The effect is that usability and user experience may be reduced when interacting with learning platforms that are not optimized for such features. Diverse design and development approaches can be adopted, but which is the better is a thorny issue. Solutions based on applications for mobile devices (native applications) exploit to the full the device's features, but they are platform specific and must be developed, updated and managed separately. Web-based solutions, on the contrary, follow the principles of open Web: they are built on technologies that are open, accessible and interoperable, and run in standards-compatible web browser [6]. As a consequence, being not developed for a specific device, web-based solutions are typically less powerful and with a poorer look and feel when accessed through a mobile device. An answer to this problem has been the use of techniques borrowed from adaptive hypermedia systems [1,8,9]. Such

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 452–457, 2014.
© Springer International Publishing Switzerland 2014

systems implement modules that modify some component of the user interface (content, layout, structure) so that it fits the user's features, environment and device. Device features include browser version and platform, availability of plug-ins, bandwidth, display size, input devices. Adaptive techniques are an effective solution but their design and implementation is costly and time consuming.

An alternative solution has arisen recently: Responsive Web Design (RWD), a paradigm that exploits new specifications for web technologies (in particular HTML5 and CSS3, not yet W3C standards) to deploy websites that respond to the user's media environment [4]. RWD enables web applications to have a better look and feel [3] and enables user interfaces to dynamically suite the user device without the implementation of specialized adaptive modules.

Given this new scenario, the question to be explored is whether RWD is able to tackle usability problems that hinder the access to learning materials through whatever device in a seamless fashion. This issue impacts on ubiquitous learning but also on openness of learning since materials and activities that are open but cannot be effectively experienced cannot be said really open.

To answer this question, we conducted an experimental evaluation aimed to test the validity of RWD vs Adaptation Techniques in a language course on an e-learning platform [5]. To perform the evaluation we switched on and off the Adaptation module and we tested the usability of the learning activity on two groups of subjects with different devices. The results show that RWD performs well with some learning activities but it does not with others which still require Adaptation Techniques to obtain a sufficient level of usability. Below we describe the instructional design of the language course and then the evaluation and results.

## 2 Designing Responsive and Adaptive Learning Activities

Stockwell [11] reports on the most salient features and issues concerning Mobile Assisted Language Learning (MALL). One of these issues actually concerns the physical characteristics of devices, a source of concern for mobile language learning [11]. Thus, MALL systems are a useful test bench for investigating the validity of RWD in learning activities since language courses are typically based on different activity types, which involve several characteristics of devices. With regard to the language course we used for the evaluation, namely a course on Collocations in English, we sketch below the main issues for its design.

*Axioms* from language and mobile learning literature.

- Language teaching assumes a variety of tasks and activities for the development of different language-related abilities [11].
- Low usability can compromise learning effectiveness and learner motivation. The issue of usability is critical to effective e-learning applications success. Research has proved that low usability can compromise learning effectiveness (see for instance [10]). Diverse reasons may be responsible for that: cognitive load increment, distraction from the main task, annoyance and sense of frustration that may even lead to abandoning the learning activity.

*Purposes* of the platform development.

- To provide teachers with a “write once run everywhere environment” so that they do not need to rewrite their learning activities for different environments.
- To develop a language learning platform suitable for both mobile and web, capable of supporting the same task and activities variety.

*Approach.*

- In order to accomplish the first purpose outlined above, we chose a web-based approach instead of native solutions. Teachers build their course through a web-based interface and the course is accessed by desktop or mobile devices through a web interface. The course is structured in progressive learning *tasks*. Each task is built by combining different base components, of two types: *Inputs* (I) and *Activities* (A). E.g. a reading task is built by associating a text (I) with an exercise (A), such as gap filling or multiple choice.
- In order to accomplish the second purpose and complying with the axioms cited above, we had to face the issue concerning the usability of tasks and activities on different devices. This issue generated the research questions below.

*Research questions*

- Are there significant differences in the perceived usability of the different tasks on different devices? Which ones seem not to possess sufficient usability to guarantee learning effectiveness?
- By using RWD procedures to adapt activities to a given device, do tasks result usable or do they not? In this case should the task be transformed using Adaptive Techniques so as to be usable on that device?

Given these research questions, we developed our project on a double track: we used RWD as guideline for delivering the learning tasks, but we also designed and implemented an Adaptive module which evaluates the degree of fitness of the learning tasks with respect to a given device (the device in use) and is able to transform the task in order to improve its usability for that device.

Table 1 shows some of the techniques we exploited to develop the language course. For each (I) and (A) component, the table shows the issues to be handled and the technologies used to face them. RWD techniques are displayed in row 2. They use a set of instructions available in the new specifications for HTML and CSS, in particular, media queries and fluid layouts. Media Queries represent the most interesting element as they allow to design style sheets so as to make them responsive to the screen and orientation of the device in use. Besides them, JavaScript libraries such as jQuery mobile, offer functions for mobile web applications. jQuery mobile implements the paradigm of progressive enhancement that allows users to access content in a layered fashion by enabling everyone to access the basic content and functionality of a web page, while also providing an enhanced version of the page to those with more advanced browser software or greater bandwidth. Problems, however, arise in case the enhanced feature is necessary and not accessory. Row 3 reports the Adaptive Techniques we implemented. The most complex ones concern learning activities (A).

**Table 1.** – Responsive and adaptive techniques used in the mobile language learning course

| | Input (I) | | | Activity (A) | | |
|---|---|---|---|---|---|---|
| | **Text** | **Audio** | **Video** | **Quiz** | **Cloze** | **Match** |
| **Issues:** *specific problems to be managed* | Legibility and readability Text length; Navigation Scrolling | Audio play-back | Video playback. Rate constraint for transmission | Variety of visualization of input display (radio and select) | Variety of keyboards with different usability and screen size | Variety of pointers with different usability; variable drag &drop support |
| **Responsive technique** *used to manage the componentes* | CSS3 and HTML5: float rule to make the text responsive to the layout; jQueryMobile to manage events of page lifecycle | jQueryMobile; HTML5 + native audio (<audio> tag) to select the supported audio format | jQueryMobile; HTML5 (<video> tag) + media query to select the supported video format and size | jQueryMobile events: tap, taphold, swipe, swipeleft, swiperight; Widget for forms | CSS3, jQueryMobile: widget for forms (text input Widget, Grid). | CSS3, HTML 5 to obtain drag & drop (draggable attribute); jQueryMobile: plugin for drag &drop and widget for forms |
| **Adaptive technique** | Text formatting in order to avoid scrolling and to increase text legibility | Player embedding. Otherwise a loss is recorded | Player embedding. Otherwise a loss is recorded | Search for most suitable task/device pair (pivot element: pointer) | Search for most suitable task/device pair (pivot element: keyboard) | Search for most suitable task/device pair pivot element: pointer + keyboard |

These techniques exploit a knowledge base representing all the relationships among language learning techniques (e.g. cloze test), language-related abilities (e.g. reading comprehension), activity types (e.g. dictation), devices features (e.g., screen size, input method) and usability problems that may arise when executing an (A) type on one device. As explained in [5], we decided to formalize this knowledge in an ontology and to define a set of rules so that a reasoner could be able to evaluate (A) fitness for a given device. In case this usability falls under a certain threshold, the Adaptive module transforms (A) in the most suited type as regards both the device's features and the ability the activity aims at exercising. In case no suited type is found, or in case the transformed activity's learning effectiveness is inferior to the original one (e.g. when a gap filling exercise is replaced by a multiple choice one, considered as less effective for a morphology exercise), the system keeps track of this loss and will later be able to compensate it with other activities.

## 3 Experimental Comparison

The aim of the evaluation was to test the hypothesis that some of the components (I) and (A) in Table 1, even adopting the Responsive Web techniques, are not usable with some devices and to test the hypothesis that transforming such components into one that the reasoner estimates to be more usable, does actually increases usability.

To test such hypotheses we used both qualitative and quantitative approaches, since each one may provide specific evidences and insights [7]. As is typical with the evaluation of adaptive systems [2] we compared two versions of the system: a SE (standard engine) version, using only RW techniques to fit content to the user device and a TE (transformation engine) version that may apply adaptive techniques when the reasoner estimates a low level of usability for a certain task on a given device.

The comparison between SE and TE was performed through an experimental evaluation using 12 real subjects, who were asked to perform two tasks on the English course on the Clire Platform [5]. The subjects were PhD students in Digital Humanities at the University of Genoa. Preliminary to the evaluation, all subjects performed an anonymous English test. All of them gained a score >75% for level A2 (European Framework of Reference for Languages). The goal of this assessment was to be sure that possible different usability results in the evaluation were not due to language difficulties of the subjects. For the evaluation, the subjects were split into two groups, homogeneous for number and for devices' features (pointer, keyboard, screen size and mobile OS). Each subject received a form which included: the instructions about the tasks, an anonymous ID for the test and a questionnaire. The Questionnaire was composed of three sections: *user data*, *device features* and *questions about the test*: a set of four points Likert-scale questions (about orientation difficulties, usability difficulties, overall rating of their experience, graphical layout) and a set of open questions about the subjects' ratings and a specific question about any alternative exercises form they would have preferred. During the tests, an *observer* annotated problems and difficulties experienced by the subjects and their execution time to complete the tasks.

Tasks were *the same* in SE and TE: two exercises on collocations, using different learning technique and different components (I) and (A). In TE, (A) could be transformed into another activity (with the same ability and learning technique) in case the reasoner evaluated it not sufficiently usable for the user device.

*Results*. The quantitative analysis showed relevant results for two parameters: *task execution time* (an indirect measure of usability) and *usability difficulties* (explicitly declared by the subjects) whose values in TE are lower than values in SE. This difference is statistically significant, as showed in Table 2, since p-values < 0.05.

**Table 2.** Usability measures with SE and TE

| | Task execution time | | Usability problems | |
|---|---|---|---|---|
| | **SE** | **TE** | **SE** | **TE** |
| **Mean** | 16,833 | 14,333 | 1,333 | 0,667 |
| **StandardDev** | 2,229 | 1,366 | 1,155 | 0,577 |
| **P-value** | **0,046** | | **0,034** | |

This overall result confirms that using adaptive techniques improves the global experience of usability. However we need a more specific analysis about the transformations made in TE in order to know which activities were transformed due to an estimation of usability lower than a threshold. Moreover, we need to know the correlation among transformations, task components and device features.

For the first goal we analyzed log files to discover which activities were transformed in TE. For the second goal, we closely examined free text data and annotations of the observers, analyzing specific tasks' components where subjects could have difficulties and we intersected these data with device features.

The most significant correlation that we discovered is between text scrolling problems in cloze tasks and medium/small screen size (r=0,94). Pointer type did not show significant differences and, surprisingly, users with Blackberry, using trackball to scroll text had less difficulties than users with touch screen.

Results from this first experimental evaluation are significant for diverse reasons. First, they show that mobile learning design has to face the issue of device variety and that the choice among different development approaches (RW, Adaptive Techniques, native solutions, etc.) has to take into account the type of learning activity. Further analyses could identify other correlations among tasks, components and device features. Second, our experimentation highlights that learning activities (i.e. exercises) suffer the problem of usability more than passive content since they require specific physical tasks in order to be accomplished (e.g., learning activities often involve heavily selecting, dragging and entering text), therefore, their design is more critical as regards mobile learning. This result is particularly relevant for language learning, as it demonstrates that usability can be gained through the integration of RWD and Adaptive techniques and an improved usability can favor a real seamless learning experience [12] (differently from current solutions of mobile learning that usually distinguish a mobile from a desktop version of a course). Finally, we can notice that RWD has anyway obtained good results with (I) components and with some activities (A) such as quiz multiple choice. Therefore, it can be argued that when the RWD framework is more widely featured by mobile devices it will be exploited by learning providers and the barriers for open and ubiquitous learning will certainly diminish.

## References

1. Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.): Adaptive Web 2007. LNCS, vol. 4321. Springer, Heidelberg (2007)
2. Chin, D.N.: Empirical evaluation of user models and user-adapted systems. User-Model, User-Adapt. Interact. 11(1–2), 181–194 (2001)
3. Jagani, S. Comparison Chart of Mobile App Development Methods. Ezine Articles. ezinearticles.com/?Comparison-Chart-for-Mobile-App-Development-Methods&id=6773577
4. Marcotte, E.: Responsive Web Design. A Book Apart (2011)
5. Mercurio, M., Torre, I., Torsani, S.: An ontology-driven transformation engine for mobile language learning. In: Proc. of the Int. Conference on Education and New Learning Technologies, Barcelona, Spain, July 1-3, pp. 2072–2082 (2013)
6. Mikkonen, T., Taivalsaari, A.: Cloud computing and its impact on mobile soft-ware development: Two roads diverged. J. Syst. Softw. 86(9), 2318–2320 (2013)
7. Nielsen, J: Risks of quantitative studies, Jakob Nielsen's Alertbox (2004)
8. Power, R., Lewis, D., O'Sullivan, D., Conlan, O., Wade, V.: A context information service using ontology-based queries. In: Proc. of Ubi-Comp, England (2004)
9. Schmohl, R., Baumgarten, U., Köthner, L.: Content Adaptation for Heterogeneous Mobile Devices using web-based Mobile Services, MoMM, vol. 230, pp. 77–86. Austrian Computer Society (2007)
10. Shneiderman, B., Plaisant, C.: Designing the user interface: Strategies for effective human-computer interaction, 4th edn. Pearson Education, Toronto (2005)
11. Stockwell, G.: Mobile-assisted language learning. In: Thomas, M., Reinders, H., Warschauer, M. (eds.) Contemporary Computer-Assisted Language Learning, pp. 201–216. Continuum Books, London (2013)

# The Learners' Expressed Values of Learning in a Media Tablet Learning Culture

Lars Norqvist, Isa Jahnke, and Andreas Olsson

Umea University
Department of Applied Educational Science
Interactive Media and Learning (IML), Sweden
{lars.norqvist,isa.jahnke,andreas.olsson}@umu.se

**Abstract.** What do learners value as learning and how does this value relate to the use of mobile technology? Inspired by Stimulated Recall and Photo Eliciting, 207 K-9 students took photos (a total of 283 photos) of what they considered learning situations. Using these photos as starting points, we conducted 21 group interviews that generated 293 utterances. From these utterances, seven themes were generated that shows how the learners expressed their values of learning in a rich and diverse way. Without asking explicitly about technology, 33% of the utterances addressed the adoption of media tablets. These statements also revealed that when the learners personally "see" a *perceived value* for them as individuals, then they label the situation as learning. This study's findings show that identifying what young learners value as learning per se should be considered when designing teaching and learning towards learner-friendly active environments for pro-sumers. These findings also open the door for developing alternative ways for evaluating projects in schools that integrate educational technology.

**Keywords:** Media tablets, Learners, Qualitative data, Photo eliciting, Schools.

## 1 Introduction

The use of media tablets in schools and higher education has rapidly increased. Several studies (e.g., [9] and [6]) illustrate the rise of media tablets in formal education. These studies most often analyse and evaluate learning from a teacher perspective, a didactical design view, or from an institutional perspective. In this study, we take the children's perspective. We argue, when education claims that there is a shift from teaching to learning [2], to learner-centred concepts, then students become pro-sumers rather than merely consumers of education. Such an approach encourages students to create and value their *own* learning experiences. If this is the case, then research should be able to study the different forms of knowledge production. In this paper, we present the learners' perspectives (K-9) with the unique approach that all pupils got media tablets. The project was launched in 2012 in a 1:1 tablet program (one tablet per child). In an earlier study [8], we focused on the teachers' digital didactical designs, which showed that the teaching practice has

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 458–463, 2014.
© Springer International Publishing Switzerland 2014

changed towards learner-centred classrooms; in this study, we explore the expressed value of learning from the learners' viewpoints. Specifically, we address two major research questions: How do children express what they value as learning? (What makes a situation to a "good" learning situation from their perspective?) (RQ1) and To what extent is the media tablet part of what the children value as learning, and if yes, how? (RQ2). Although the main purpose of the study was to explore how learners do experience the use of media tablets, we did not want to focus on the media tablet itself. Instead, we focused on learning to see whether the learners related learning to the media tablet or not.

## 2 Description of the Methods and Context

This empirical study was conducted in a municipality in a Scandinavian country with approximately 22,000 inhabitants. The municipality has seven schools from preschool class to grade 9 with around 170 teachers and 2,000 pupils aged 6-16. In January 2012, all teachers and students were given a media tablet (iPad). This study is part of a larger project that evaluates the integration of media tablets in all schools in the municipality. This project has become a project of "changing the teaching and learning culture" with the aim to develop a learner-friendly environment in which the children are active and reflective pro-sumers making their learning visible [8].

The study design was presented in meetings with teachers and school leaders. These meetings resulted in 24 teachers expressing an interest in participating; voluntary sampling [3]. In the end, 11 teachers and 11 classes participated. We requested the learners to take a photo of a "good" learning situation and then mark it with one word. The situations could be situated in school or in any other context. Thus, the learners' perspective includes both a) the school context and b) the non-school context. A child doesn´t have the primary identity of a student or pupil outside of school but can be seen as a learner or child in both school and non-school learning settings, therefore we use the term *learners* or *children.*

The data were collected using specific interview techniques. During the interviews, the children were asked to identify "good" learning situations as depicted in photographs they had taken (phase 1). We also led a discussion about what they think "good" means in the specific context (phase 2). The photographs, taken by the children and mainly with the media tablet, were marked with a word or a short sentence provided by the children. This approach is inspired by Stimulated Recall [7] and Photo Eliciting [4]. The photographs and the utterances from the children were used to understand what the children believed learning is in general and in relation to the use of media tablets. Phase 2 consisted of mixing forms of semi-structured interview methods [5] [10] such as e*xplorative conversation*, *focus group interviews*, [12] and p*hoto-interviews*. The interviews used open ended questions such as "What does this photograph show?" or "Why did you pick this photograph?" Usually, two audio recording devices were used to record the conversations and interviews.

The interview guide focused on learning situations and not on technology. In the first phase, we did not ask about the relation of learning to the media tablets, as we did not want the children to focus on media tablets but rather on learning. We assume this approach strengthens the validity of the data that describe the actual adoption of mobile technology in classrooms.

This study analysed what the children said about the photographs they took. All the utterances were collected in themes using a software programme. Seven themes were coded from the total sample of utterances. In total, the study involved 11 classrooms from grade 1 to grade 7. There were 207 children (100 girls and 107 boys) and 11 teachers (9 female and 2 male). The children produced 278 photographs and five drawings. A total of 293 utterances were generated.

## 3 Findings

The findings are organized according to the seven themes that were generated from the analysis of the 293 utterances. The seven themes show the children's view on "what learning is" with respect to their definitions of learning situations (Table 1):

1. Various locations;
2. Reading and practising language skills;
3. Learning from someone (traditionally face-to-face);
4. Learning from/through something (mediated artefacts);
5. The learners know what works for them –'I did it my way';
6. Engaged interactions;
7. Making mistakes is OK.

| **Utterances** | **Themes** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **Total** |
| Single meaning utterances (SiM) | 44 | 52 | 26 | 50 | 56 | 48 | 17 | **293** |
| Not related to Media tablets | 30 | 34 | 23 | 12 | 46 | 38 | 14 | 197 |
| Related to media tablets | 14 | 18 | 3 | 38 | 10 | 10 | 3 | 96 |

**Fig. 1.** Summary of utterances in seven themes including relation to media tablet without explicitly asking about the media tablets.

Theme 1 was generated from utterances that expressed learning activities in different locations: e.g., "*Homework – To get the homework from the Intra*"; "*BookCreator – Rather work alone, but it can be done with others*"; and "*Bake a cake*". We also asked follow-up questions such as "What do you learn when you bake?" One girl responded that she mainly learns about math when baking (e.g., decilitres and litres). The first theme was generated from 44 utterances; 14 utterances were related to the media tablet. In total, 32% of the utterances were related to the media tablet even though the interviewers never explicitly asked about technology.

Theme 2 was generated from utterances about reading and practising language skills (i.e., writing or spelling assignments). From the children's view, this type of learning was seen as a prerequisite for more learning and enrichment of the language: "*To have a vocabulary – The meaning of word is important*"; "*To read is most* *important, you need it/do it all the time*"; "*Learn by reading – you get a better education by reading"*. The second theme was generated from 52 utterances of which 18 (35%) were related to the media tablet.

Theme 3 was generated from 26 utterances that related to directly receiving help, information, or guidance from a teacher, a peer, or a parent: e.g., "*You can learn from someone else who knows more*"; "When I as a learner can observe and hear another person do something"; "*From someone who is explaining – It can be an adult or a peer*"; "*When the teacher explains – in his/her own language. It´s the best when the teacher explains on the blackboard since it also is written and can be seen*". Three (12%) of these utterances referred to the media tablet.

Theme 4 was generated from 50 utterances about receiving help, information, or guidance from *something* (i.e., no direct/live interaction or conversation) that supports learning. Obviously, there is someone "behind" these things, but theme 4 focuses on the fact that learning has taken place using mediating artefacts. In most cases, the children's utterances referred to a media tablet and the features: "*You first see a YouTube film about something and then you practice*"; "*Google – you can find out something that you don't know. For example, if you need help with the computer – the whole world can help you*"; "*Apps – Different apps to learn different things*". Of these utterances, 38 (76%) were related to the media tablet.

Theme 5, 'I did it my way', was generated from 56 utterances that illustrate individual ways of learning and a deeper level of individuality. These learners valued learning when they knew what works for them during the learning process in order to improve the learning progress: "*Find out new things*"; "*To explore new topics*"; "*A school visit to a piggery - Learning is facilitated when something feels real, to learn when it is for real*"; "*Practise – sit at home and practice alone*". Of these utterances, 10 (18%) were related to the media tablet.

Theme 6, engaged interactions, was generated from 48 utterances expressing a form of communication where the learners interacted, collaborated, and received help from someone. It also included expressions about sharing or presenting something: "*It is important to be together in the class because we all have different views*"; "*Group work – You learn through discussion. Depending on the subject, group work can be more or less common*"; "*When you show/present things, it makes you repeat, write, prepare. Then you learn*". Of these utterances, 10 (21%) related to the media tablet.

Theme 7, "*making mistakes is OK*", was generated from 17 utterances about learning activities such as initiatives and challenges. In this theme, the utterances express learning as trials in which the children learn from being wrong or making mistakes: e.g., "*At first it didn´t work but then it worked*"; "*Learn from being wrong; to make it better another time. It´s ok to be wrong in school*"; and "*Try on your own*". Of these utterances, 3 (18%) were related to the media tablet.

## 4 Discussion and Conclusion

The learners expressed their values of learning in a rich and diverse way (RQ1). The children valued learning when their personal needs were addressed (e.g., when something feels real) and when they saw benefits for themselves (theme 5, n=56 utterances). The children stressed different values of learning as individuals and valued learning in situations where they learned from/through something (theme 4, n=50), in reading and practising language skills (theme 2, n=52), in engaged interactions (theme 6, n=48) (theme 1, n=44), when they learned from someone in traditional face-to-face communication (theme 3, n=26), and when making mistakes was not stigmatised (theme 7, n=17). In other words, the learners preferred learning when they saw a benefit from expanding their pre-existing knowledge. The findings show that a learner-centred classroom takes advantage of students' perceived values of learning when designing teaching and learning.

Of the 293 utterances collected, 96 (33%) were related to the media tablet (Fig. 1) (RQ2). The utterances that are summarized in the table came from the children (open coded) when asked to describe learning from their point of view (see Methods). When around one third of the utterances have a connection to the tablet, without explicitly asking for the use of it, then, we argue, this is a *strong indicator* that the media tablet became natural part to the school culture; it is like a book or another resource for learning. When the schools would remove the tablets, then one third of the learning situations suddenly were lost. After a 2 year of using the media tablets everyday in the classroom, the tablets were not perceived as being special anymore or as something the teachers needed to encourage the learners to use as the learners used the media tablets when they needed. Depending on the learning activity, the teachers and learners decide whether the tablet is helpful to use or not. The media tablet can be seen as integrated naturally in the teaching and learning setting. However, the quotations revealed that a media tablet is not used all the time. The learners had already learned that sometimes the tablet was not useful. The study shows that it takes more than a media tablet to support the learners' perspectives of learning [11].

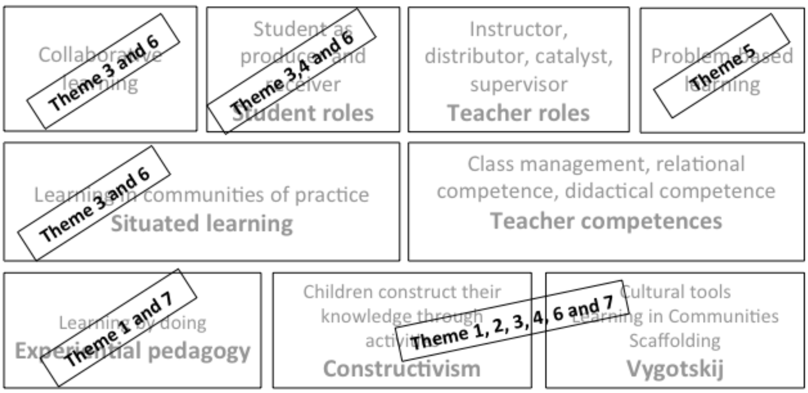



**Fig. 2.** The municipality framework in the schools that participated in our study. It can be used to evaluate their teaching and learning practices via connecting the themes.

**An Alternative Approach for Evaluating Education.** With an understanding of the learners' view of "what learning is", the municipality can evaluate their teaching and learning practices. The municipality that participated in our study used a model to develop and evaluate the schools' understanding of teaching and learning (Figure 2). From such a framework, the seven themes presented in this article can be analysed and connected to the different blocks in the framework, resulting in an evaluation of the schools' teaching, learning, and education activities. When applying the themes from the learners' view of "what learning is" to the municipality project framework, developing teaching and learning with respect to mobile technology can be evaluated from the perspective of what the schools value. This strategy gives the schools an alternative method of evaluating teaching, learning, and the use of mobile technology. Schools will not have to rely only on national or international instruments such PISA.

**Conclusion.** This qualitative study explored how young learners express *what they value as learning* in a municipality with a 1-1 media tablet programme. Photos that depicted learning situations were used as starting point. Media tablets are integrated in the learning culture and the relation between the media tablets and what the learners value as learning is expressed in a third of all utterances. The innovative methodology of this study reveals the learners perceived values of learning in a media tablet learning culture. This study also points to a new question: How to design teaching and learning that it increases the learners' perceived value of learning? Using the data from the teachers' view and classroom observation [8], we hope to develop a more comprehensive picture of the change in schools from teacher-centric to learner-centred classrooms in future studies.

## References

1. Attard, A., Di Loio, E., Geven, K., Santa, R.: Student centered learning: An insight into theory and practice. In: Partos Timisoara, Bucharest, pp. 6–15 (2010)
2. Barr, R.B., Tagg, J.: From teaching to learning—A new paradigm for undergraduate education. Change: The Magazine of Higher Learning 27(6), 12–26 (1996)
3. Bryman, A.: Social research methods, 3rd edn. Oxford University Press, Oxford (2008)
4. Cappello, M.: Photo Interviews: Eliciting Data through Conversations (2005)
5. Cohen, L., Manion, L., Morrisson, K.: Research Methods in Education. Routledge, London (2007)
6. Findahl, O.: Svenskarna och internet 2013: SE (Stiftelsen för Internetinfrastruktur) (2013)
7. Haglund, B.: Stimulated recall: Några anteckningar om en metod att generera data. Pedagogisk forskning i Sverige 8(3), 145 (2003)
8. Jahnke, I., Norqvist, L., Olsson, A.: Teachers' Digital Didactical Designs of Learning Expeditions. In: Proceedings of the ECTEL 2014, Graz, Austria (2014)
9. Kaganer, E., Giordano, G.A., Brion, S., Tortoriello, M.: Media tablets for mobile learning. Communications of the ACM 56(11), 68–75 (2013)
10. Kvale, S., Brinkmann, S.: Interviews: Learning the craft of qualitative research interviewing. Sage (2009)
11. McCombs, B.L.: Assessing the Role of Educational Technology in the Teaching and Learning Process: A Learner-Centered Perspective (2000)
12. Patton, M.Q.: Qualitative research & evaluation methods. SAGE, London (2002)

# Can Teachers' Fragmented Work Situation Jeopardize Professional Development of Future Teaching Practices?

Lena Pareto and Sara Willermark

University West, Department of Economy and IT,
461 86, Trollhättan, Sweden
{lena.pareto,sara.willermark}@hv.se

**Abstract.** Digitization of schools has been on the political agenda for decades, but despite all reforms, technology investments and professional development initiatives, the goals are not yet reached. We have examined how teachers' perceive their working situation in order to explore if the conditions are suitable for learning and novel teaching practices. 18 interviews with Swedish primary school teachers were conducted and transcribed, from which 330 excerpts were extracted and divided in two characterizing categories: *fragmentation* reflecting working rhythm and *density* reflecting working tempo. The working condition had characteristics known to cause stress and less wellbeing, which counteracts teachers' sensitiveness to adapt to novelties and a reflective practice. This may jeopardize future professional development and thus digitalization of schools.

**Keywords:** digitalization, school, 21st century skills, professional development.

## 1 Introduction

Government reports all over express the need to meet the demands of the 21st century skills, where emphasis is made on problem solving, teamwork, peaceful conflict resolution, dealing with complexity and ambiguity, and being life-long learners [1]. The European Parliament listed digital competence as one of eight key competencies for the future. Digital competencies are nowadays referred to as "life skills", comparable to literacy and numeracy, thus becoming "both a requirement and a right" [2].

The digitization of schools has been on the agenda for decades and has been intensified in recent years. Policy makers, technology producers and educators all express different hopes and expectations concerning what such implementation can lead to (e.g. [3,4]). There are numerous initiatives in European countries to integrate ICT in education. Extensive investments have resulted in Swedish schools being equipped with technology above average whereas the computer usage is average compared to European countries [5]. Nonetheless ICT equipment in the classrooms does not per se create new teaching practices aligned with 21st century skills. Many teachers and pedagogical ICT-applications use the new medium in a substitutive manner by "reinforcing old ways of teaching and learning" [6]. It is often claimed that the benefits of ICT-based teaching lies in transforming learning into new forms and redefining the nature of teaching, e.g., the TPACK framework [7] which suggests that teachers need

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 464–469, 2014.
© Springer International Publishing Switzerland 2014

expertise in technology, pedagogy, content and context as interdependent qualities necessary to teach effectively with technology. Others also claim that the new competencies demanded by society call for radical change [1][8]. Yet, teacher training and education reforms have traditionally been quite divorced processes [1]. Research has shown a great complexity related to school reforms [9,10] and implementation of technology is no exception [11]. The discrepancy between policy and reality in terms of implementation of ICT is often blamed on the schools [12].

This study is a detour from a 2-year teacher-training program financed by the National School Board in Sweden, in which an innovative pedagogical game [13] was used to explore novel teaching practices in the classroom. The training followed advocated methods where ICT-usage is integrated as a natural ingredient in the didactic process [14], supported by known success factors such as collegial collaboration, classroom observations, outside expertise, reflection and experimentation [15, 16]. Yet, highly dedicated and skilful teachers felt forced to limit their participation and engagement beyond their desire, which intrigued us to investigate why.

The research question we address concerns whether *the working condition for teachers allows for radical change such as technology-inspired redefinition of teaching practices?* We will search for impeding factors in the working situation that may counteract the desired change due to the digitalization of school and society.

### 1.1 Related Work of Teachers' Working Situation and Interruptions at Work

The issue of teachers' working situation has received attention: OECD is currently conducting an international study, which comprises identification of teacher's time distribution across countries. A recent report from the Swedish National School board shows that Swedish teachers work 9 hours and 40 minutes per workday, as well as on weekends. In addition to long working hours and little opportunity for rest; teachers' working situation is characterized as being intense and complex [8][17]. In contrast to many professions who encounter their clients individually, teachers mainly meet their students in a collective situation but are still expected to individualize their teaching, causing a tension between individual needs and the group's need. During non-teaching hours, all other tasks such as administration, documentation, planning, and assessment, and lately new duties such as collaborations with other professions, mentoring, and numerous school development programs, need to be accomplished. All together this leads to working days full of activity switches and interruptions.

The performance and psychological effects of interruptions at work was examined in an experimental study [18]. The experiment participants carried out a realistic office task, while the frequency and complexity of interruptions were experimentally manipulated. Unexpectedly the study showed that interruptions cause people to perform the task *faster* while maintaining the level of quality: the gain due to over-compensation is larger than the interruption cost. However, interruptions have negative impact on emotions and wellbeing, and lead to higher effort levels, so the improved performance is achieved at the expense of higher psychological costs.

A similar subsequent study [19] supports the previous findings that people compensate for interruptions by working faster, at the cost of more stress, higher frustration, more time pressure and effort. In addition they found that the interruption cost was independent of the topic being related or unrelated to the interrupted activity.

# 2 The Study

The importance of studying teachers' work conditions sprang from a national program on developing mathematics in schools using ICT. The project was conducted an action-based research collaboration between researchers and practitioners, and involved in total about 60 teachers and 1000 students from a municipality in West Sweden.

To address the research question, we chose to investigate the teachers' perceptions of their work conditions by means of interviews in an explorative, interpretative approach. The respondents were recruited from participants in our project. The study includes 18 interviews with teachers, representing ten different schools and areas of socio-economic environments.

The interviews were based on an open conversation in dialogue with the respondents, using a set of open-ended questions for the interviewer to choose from; all with the intention to get the respondents in a narrative frame of mind to describe and give examples based on their own experiences. The interviews were conducted at the respondent's respective workplace and lasted between 40 to 90 minutes. The interviews were recorded and transcribed. During the interviews and transcription process, the idea of fragmentation recurred in the respondents' stories and appeared as a central and commonly experienced phenomena worth analysing further. The transcripts were analysed in a grounded theory spirit, and two researchers extracted excerpts related to time and fragmentation issues.

# 3 Results

From the interview transcripts, there were 330 excerpts judged as related to time- and fragmentation issues. These were divided in two main categories, describing the teachers work situation; *fragmentation* (76 excerpts) and *density* (206 excerpts), and one category related to *perceived consequences* of these issues (48 excerpts).

## 3.1 Fragmentation

By fragmentation we refer to causes or events behind activity switches during work, which together constitute *the working day rhythm.* A total of 76 excerpts of teachers' work situation can be related to fragmentation, where teachers' work is characterized by interruptions and lack of continuity. 25 excerpts related to fragmentation due to interruptions resulting in disturbances, during teaching or other tasks. The interruptions were for example caused by student conflicts, phone calls from parents or unplanned meetings with colleagues: "*There is always someone who interrupts you, if you try to sit down for 10 minutes to reflect*".

There were 51 excerpts related to 'Lack of continuity', which refers to the agility of teachers' profession and how they frequently have to handle new demands, rapid changes and unforeseen events. Extracts related to keeping up with new knowledge, new teaching methods, tools and techniques, is exemplified by: "*The profession is constantly changing, there are new teaching tools and new research which generate new educational ideals*" or "*All the new technology, I find it hard to allocate time to keep up with technological development*". Furthermore, respondents describe how they must handle many activity switches during a workday and also deviate from their original plan due to changed circumstances, such as substituting for absent colleagues, or modifying their original plan due to technology failure.

### 3.2 Density

Density refers to causes behind the total amount of activities to be handled during work, which determines how fast tasks need to be accomplished and how this give rise to the *working day tempo*. A total of 206 excerpts of teachers' work situation can be related to density, which denotes quotes that refer to how full of activities a working day is, and reasons for that. The respondents express how intense their preplanned schedules already are, and that pauses and time to reflect is not achievable in practice. Respondents express how they must handle students varying needs as well as handle a multitude of roles, related to the teaching profession. It describes the variety of functions and tasks carried out by teachers, both professional activities, but also additional tasks such as guard the schoolyard or serve lunch. A total of 109 excerpts were categorized as different types of "time absorbents" and comprise work assignments that are perceived as (too) time consuming. Examples include collaboration with colleagues, behaviour exhortation or documentation leading to a reduction of time control. In addition to technical problems and non-functioning equipment, learning and implementing new technology were considered as time absorbents as well, as illustrated by: "*All the new computer programs and all the opportunities available, but when will I find the time to learn more and update myself?*"

### 3.3 Perceived Consequences

The category includes 48 excerpts and refers to statements expressing teachers' perceived consequences of their work situation: deterioration due to shortage of time, overwork and stress. The respondents express how the pressed work situation hinders them from planning innovative high-quality lessons and how the pedagogical issue, as a consequence, is given lower priority "*The stress makes me less of a teacher*". Additionally respondents express a need to overwork, in order to keep up with their duties and how they perceive stress due to their work situation, as illustrated by: "*At some point it simply becomes too much to handle that the school finally had to send me to counselling*".

## 4 Discussion

The respondents provide a consistent view of the profession as fragmented and intense with an increasing workload. There are mainly two ways to handle the situation: (1) To work more: this study indicates that teachers overwork and the result is consistent to previous studies. (2) To take time from other assignments: the respondents expressed that the time for planning and reflection was often consumed by more immediate, insistent tasks difficult to ignore. The respondents expressed a great frustration over their work situation. Yet they do not indicate much resistance towards their working conditions, instead they find strategies trying to cope with the situation.

Thus, they compensate for the high workload and the fragmented work situation by working faster and harder. This is aligned with the results in [18] and [19] which shows that interruptions cause people to put in more effort and perform tasks faster by developing strategies to deal effectively with the interruptions. Already in the short-term perspective, this caused psychological costs in terms of more stress and a negative impact on emotion and wellbeing. In addition, our respondents also expressed, that they were not able to maintain the same quality of their work any longer. This is likely an additional long-term effect of a fragmented and dense work-situation.

Based on these results, we would argue that the teachers' working situation is far from optimal and that a reflective practice is hard to achieve under the current strain of work intensification. Therefore, the requested change on the teachers' behalf to redefine teaching practices better aligned with 21$^{st}$ century skills development and the digitalization of society is jeopardized by the current work situation. Thus, teachers' ability to make a difference to education and to society is at stake [8].

## 5 Conclusion

As a result of our investigation, we have identified the following characteristics of the working situation: a fragmentation of time, fragmentation of responsibilities and density of activities. These factors decrease the teachers' opportunity to deploy a reflective practice, to engage in transformative learning and to develop novel teaching practices. Hence, the eagerness to accelerate the digitalization of schools by providing more professional development initiatives, more reforms and demands for change, may not only have little effect it may be contra-productive. As a consequence, the fragmented and intense work situation can jeopardize teachers' future professional development and thus a productive digitalization of schools.

## References

1. Moreno, J.: Do the initial and the continuous teachers' professional development sufficiently prepare teachers to understand and cope with the complexities of today and tomorrow's education? Journal of Educational Change 8, 169–173 (2007)
2. Ferrari, A., Punie, Y., Redecker, C.: Understanding digital competence in the 21st century: An analysis of current frameworks. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 79–92. Springer, Heidelberg (2012)

3. Hurd, S.: Why has computer assisted learning made so little impact in secondary education? Lessons from an economics and business subject case-study. The Curriculum Journal 20(2), 139–159 (2009)
4. Ward, L., Parr, M.: Revisiting and reframing use: Implications for the integration of ICT. Computers & Education 54, 113–122 (2010)
5. European Schoolnet Survey of Schools: ICT in Education. Benchmarking Access,Use and Attitudes to Technology in Europe's Schools (March 22, 2014), http://ec.europa.eu/digital-agenda/sites/digital-agenda/files/KK-31-13-401-EN-N.pdf
6. Resnick, M.: Sowing the seeds for a more creative society. Learning and Leading with Technology 35(4), 18–22 (2007)
7. Mishra, P., Koehler, M.J.: Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. Teachers Collage Record, 108(6) (2006)
8. Bottery, M.: Education and globalization: redefining the role of the educational professional 1. Educational Review 58(1), 95–113 (2006)
9. Bryk, A.S., Bender Sebring, P., Allensworth, E., Luppescu, S., Easton, J.Q.: Organizing schools for improvement. Lessons from Chicago. The University of Chicago Press, Chicago (2010)
10. Elmore, R.F.: School reform from the inside out. Policy, practice and performance. Harvard Education Press, Cambridge (2004)
11. Erstad, O.: Digitalt kompetente skoler. In: Erstad, O., Hauge, T.E. (eds.) Skoleutvikling og Digitale Medier. Kompleksitet, Mangfold og Ekspansiv læring (School development and digital media: complexity, diversity, and expansive learning), pp. 47–63. Gyldendal Akademisk, Oslo (2011)
12. Olofsson, A.D., Lindberg, J.O., Fransson, G., Hauge, T.E.: Uptake and Use of Digital Technologies in Primary and Secondary Schools – a Tematic Review of Research. Nordic Journal of Digital Literacy 6(4), 207–225 (2011)
13. Pareto, L.: A Teachable Agent Game Engaging Primary School Children to Learn Arithmetic Concepts and Reasoning. International Journal of Artificial Intelligence in Education 24(3), 251–283 (2014)
14. Hameyer, U.: School improvement as a problem solving process: Using what we know. In: Letschert, J. (ed.) Turning the Perspective: New Outlooks for Education. CIDREE/SLO, Enschede (2001)
15. Hattie, J.A.C.: Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge, London (2008)
16. Wilson, G., Stacey, E.: Online interaction impacts on learning: Teaching the teachers to teach online. Australasian Journal of Educational Technology 20(1), 33–48 (2004)
17. Eraut, M.: Menus for choosy diners. Teachers and Teaching: Theory and Practice 8(3/4), 371–379 (2002)
18. Zijlstra, F.R.H., Roe, R.A., Leonora, A.B., Krediet, I.: Temporal factors in mental work: Effects of interrupted activities. Journal of Occupational and Organizational Psychology 72, 163–185 (1999)
19. Mark, G., Gudith, D., Klocke, U.: The cost of interrupted work: more speed and stress. In: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems, CHI 2008, pp. 107–110. ACM, New York (2008)

# An Augmented Reality Job Performance Aid for Kinaesthetic Learning in Manufacturing Work Places


Fridolin Wild[1], Peter Scott[1], Jaakko Karjalainen[2], Kaj Helin[2],
Salla Lind-Kohvakka[3], and Ambjoern Naeve[4]

[1] The Open University, Knowledge Media Institute, Milton Keynes, UK
{fridolin.wild,peter.scott}@open.ac.uk
[2] VTT Technical Research Centre of Finland,
Human Factors in Complex Systems, Tampere, Finland
{jaakko.karjalainen,kaj.helin}@vtt.fi
[3] If P&C Insurance, Niittyportti 4, Postbox 1032, FI-02200 Espoo
salla.lind-kohvakka@if.fi
[4] Royal Institute of Technology (KTH), Stockholm, Sweden
amb@csc.kth.se



**Abstract.** Manufacturing – as the name already suggests – is an area where novel 'learning by doing' approaches can help tackle the skills gap this industry is currently facing at large. Within this contribution, we apply human-centred design methodology to help overcome two significant 'teething problems' for augmented reality. We propose five use cases for the application of learning technology on- and off-the-job and we show how these use cases can be implemented in an augmented reality based job performance aid, which we subsequently test on a real life story board in the furniture industry.



**Keywords:** Augmented Reality, Job Performance Aid, Kinaesthetic Learning.


## 1 Introduction

The third generation of commercial augmented reality (AR) applications and devices is being heralded (e.g. [1]) and – as with all novel technology – it is not clear whether it will become mainstream or only cater niche markets. Other than its two predecessor generations, this third wave not only provides location and viewpoint adaptation, not only links content through computer vision to physical objects (ditto), but promises improvements in human computer interaction through wearable devices and a generally more immersive "unity of machine and human, working towards their common goal", as [2] put it in their recommendation to knowledge management and learning technology strategists in industry on how to develop virtual and augmented reality systems that help leverage human capacity for performance at the workplace.

The deciding factor whether this third wave of AR-software is mature enough for a mass market, so[1], is whether its four main teething problems will be addressed: fix the lack of critical content (incl. user-generated and repurposing of web contents), innovate better use cases, enable 'fusion tracking' of reality registered to a single


C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 470–475, 2014.


© Springer International Publishing Switzerland 2014

reference space, and create more optimal and accepted presentation techniques. While it may not be possible (or desirable) to find generic answers for these problems, the niche of learning technology in the manufacturing industry may be confined enough to enable development of viable solutions for on- and off-the-job learning with the help of performance aids.

Augmented reality based technology enhanced learning solutions and job performance aids for manufacturing workplaces may well be a potential solution on which to demonstrate that technological teething problems can be overcome, while at the same time supporting an industry in its significant transformation and – if not prevent – then at least soften the expected (further) blows to the labour market by scaffolding personnel in their skill development to the next level.

To make this happen and to tackle the four problems mentioned above, we have already provided an initial *single reference space* in previous publications [3,4], which we call 'workplace and workspace model'. This reference space enables fusion tracking across sensors and devices by registering them in a unified workplace model, an instantiation of a more general workspace model of the manufacturing industries. Moreover, we have started integrating *content learning object repositories* for manufacturing industries in furniture production, aviation maintenance, and textile production, with tens of thousands of content objects available on the intranets at our industrial partners Agusta Westland, Aidima, and DITF (see [4]). Additionally, we have set up social learning platforms[1] for these industries that will help elicit and share user-generated content.

With this publication, we report on how we tackle the two remaining teething problems for augmented reality based job performance aids, namely the problem of lacking *use cases* and the problem of more *optimal presentation* technology.

The remainder of this paper is therefore organised as follows. In the second section, we tend to the development of *novel use cases*: we derive a set of five fundamental application cases that implement required bi-directional tacit-explicit knowledge conversion modalities. These use cases help to determine the right balance between on-the-job and off-the-job elements for the different usage contexts. Following in Section 3, we then tend to the *improved presentation technologies*: we report concept and architecture of the presentation system, to – in Section 4 – subject the prototype to a reality test by adapting it to a full storyboard (in artisan furniture production), testing the use cases introduced. An outlook wraps up the paper.

## 2 Use Cases in Manufacturing Learning

There are conceptual differences in how learning is best supported in different application contexts given in the workplaces studied (in furniture production, textile production, and helicopter maintenance), whether those needs are away from the workplace – e.g., in 'training' mode somewhere else – or the needs are at the workplace – e.g., in 'performance-support' mode. The key difference is between on-the-job

[1] TELLME .live: `http://tellme.kmi.open.ac.uk`

'learning while doing' performance-support and work-interrupting training – and below we will further differentiate on-the-job and off-the-job into five distinct use cases. Consequently, as this impacts on the cognitive and social capacity available, the mix of modalities for interaction and delivery for each use case has to differ, catering for the appropriate level of attention workers can direct to presentation and interaction with content. It is useful to differentiate between traditional (web-based) learning content and traditional augmented reality content.

**Table 1.** Main modalities in tacit-explicit knowledge conversion of the envisioned application scenarios

| | Tacit | Explicit |
|---|---|---|
| Tacit | *Socialisation*: Debriefing, On-the-job | *Externalisation*: Wicked Problem, Debriefing |
| Explicit | *Internalisation*: Briefing, On-the-job, Wicked Problem | *Combination*: Just about to start working, Wicked Problem |

Traditional learning content typically entails text and multimedia files presented in a web browser user interface. It sometimes encompasses more complex web apps and web widgets providing access to interactive content. Traditional learning content can be thought of as digitised classroom learning material. Traditional augmented reality (AR) content can be defined in the words of Radkowski & Oliver (2014), who highlight the following main characteristics: "superimposes the natural visual perception of a user, generally provided by a video camera, with computer-generated information". AR enables the contextualisation of learning content from its linearity (e.g. long text paragraphs) to space (e.g. using action symbols by tying content to its environmental context). The five use cases derived target different temporal stages in workplace learning and they differ in the mix of traditional learning content versus AR-based content.

*Briefing* is typically done in meeting-room or classroom-like environments. Users receive the big picture by going through written descriptions of the different steps of the process and all required background information. This scenario is optimal for traditional learning content.

*Just about to start working*: The user is already in the place where actual work will be conducted and the environment context can be utilized to enable AR content. It is optimal to deliver a mix of traditional learning content with multimedia elements, while binding it to the environment with the help of AR technology. This kind of mix can be implemented on a mobile device using, e.g., a tablet PC or glasses.

*On-the-job*: When actual work has started, the cognitive capacity of workers can be severely limited. This is where multimodal AR-technology is at its best.

*Wicked problem* situations emerge in actual on-the-job contexts, leading to interruption or even abortion of the current work task in process. In this use case it is vital to combine the strengths of context-awareness provided by the AR component together with all the possibilities of traditional content presentation.

*Debriefing*: When learning work has finished, it is time to recapture what was learned, also involving authoring, for example, in form of evidence capturing.

These five fundamental use cases – briefing, just about to start working, on-the-job, wicked problem, debriefing – relate to the four archetypes of tacit-explicit knowledge conversion required for workplace learning as depicted in Table 4: for example, tacit-tacit conversion ('socialisation') is facilitated by the debriefing and on-the-job scenarios.

## 3 Architecture and Components

To implement the use cases, special emphasis was given in the prototype to enabling a bi-directional link between the traditional augmented reality user interface and the web-app HTML5 components characteristic of traditional learning content.

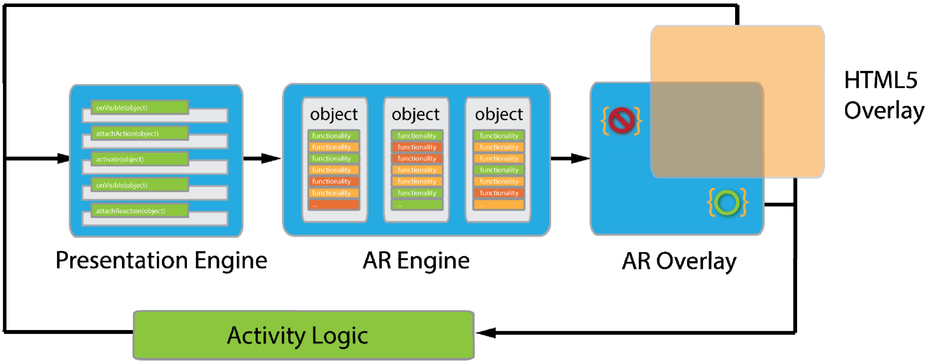



**Fig. 1.** The components

The presentation engine is used to control the AR objects provided by AR engine (see Fig. 3). It can enable and disable AR components depending on the current activity context. For example, it is possible to activate markers relevant in the current context, so that the AR engine restricts detection to what is currently relevant. Moreover, it is also possible to enable and disable functionality of individual AR components, making it possible to e.g. provide audio and visual cues in some occasions and visual only in others, while still using the same basic AR component.

Fig. 4 represents the basic viewport layout of the user interface. The HTML5 overlay features three viewports, from left to right: the location and tinker belt; the actions pane; the reactions pane. The location thereby lists (once detected by the underlying AR layer) the current location. The tinker belt is used to list tools detected in the real world that are relevant for the current learning activity. For example, if an on-the-job learning activity requires a spray gun to be used, it will show up with a graphical representation here, once it has been detected.

Actions and reactions encapsulate the instructional user interface side: actions relate to instructions, e.g., on what tools to pick up or what to do with them. Dialogue messages can be inserted here that require certain ‘reactions’: if a micro-assessment is conducted by posing a question, the HTML5 widgets representing potential choices will be listed in the reactions pane. An example of this is e.g. a multiple-choice question about work safety – listing correct and wrong answers in the reactions viewport.

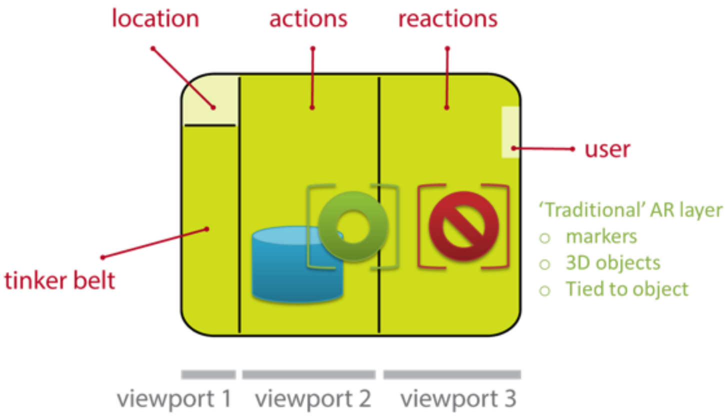



**Fig. 2.** Basic model for AR enhanced tablet UI with HTML5 overlay


**Fig. 3.** Snapshot from the demo of the implemented storyboard

The data required for presentation and summarisation of the storyboard as a learning experience is encoded in a set of XML files, that can be packaged with the actual app or that can be fetched live from a web service. The XML files hold representations of the locations, actions, reactions, users, and objects. The app is implemented in Unity3D, using CoherentUI.

## 4 Testing the Solution on an Industrial Case

Fig. 3 shows a snapshot of the user interaction in the implemented storyboard. After the user checks in (using a badge to authenticate), interaction with different materials and machines is demonstrated (sanding machine/papers with varying grit size in the sanding box; spray gun and different-sized nozzles in the spray booth). The prototype tests, how micro-assessments can be integrated to support data collection and matching beyond automated observation of the visible by posing simple multiple-choice questions in the existing dialogue structuring facilities available. The prototype demonstrates complex object interaction, guiding the user through a flow of actions, thereby requiring context-sensitive activation and deactivation of markers.

## 5 Conclusion and Outlook

With the development and the prototype for tablet PCs and glasses, we already have discovered that interaction between components will pose an additional challenge and the integration of widget components directly into an overlay HTML5 layer is not desirable at all times (particularly for the glasses application).

We therefore envision for future development a stronger role of the widgets and interactive eBooks accessible on the tablet and other mobile devices and we will work towards stronger interlinking the interaction between them. For example, an eBook with embedded video material and interactive HTML-based widgets can be used to initiate interaction in the glasses (think of a 'try this activity' widget that triggers the augmented learning experience to be started in the glasses).

Testing of the first integrated system will provide additional valuable information regarding usability and utility of the first version. As the technology will be directed to serve a range of different purposes and industries, the development work may include several compromises to emphasize the particularly strong and rewarding features. System development will continue exploiting human-centred design methods and its iteration loops with end-users from industry.

## References

1. Langlotz, T., Grubert, J., Grasset, R.: Augmented Reality Browsers: Essential Products or Only Gadgets. In: CACM, vol. 56(11) (2013)
2. Krassi, B., Kiviranta, S.: The unity of human and machine. In: VTT Impulse: Science, vol. (1), pp. 24–29 (2013)
3. Naeve, A., Isaksson, E., Lefrere, P., Wild, F., Tobiasson, H., Walldius, A., Lantz, A., Viitaniemi, J., Karjalainen, J., Helin, K., Nuñes, M.J., Martín, J.: Integrated industrial workplace model reference implementation, deliverable d4.3. TELL-ME Consortium (2014)
4. Cattaneo, F., Megliola, M., Sesana, M., Wild, F., Scott, P., Isaksson, E., Naeve, A., Maschler, T.: Distributed Learning Repositories Access Methods, deliverable d3.1. TELL-ME Consortium 1 (2013)
5. Radkowski, R., Oliver, J.: Enhanced natural visual perception for augmented reality-workstations by simulation of perspective. Journal of Display Technology 10(5) (2014)

# Have You Found the Error? A Formal Framework for Learning Game Verification

Amel Yessad, Isabelle Mounier, Jean-Marc Labat, Fabrice Kordon, and Thibault Carron

LIP6, CNRS UMR 7606, Université P. & M. Curie – Paris 6
4, place Jussieu, F-75252 Paris Cedex 05, France
{Amel.Yessad,Isabelle.Mounier,Jean-Marc.Labat}@lip6.fr,
{Fabrice.Kordon,Thibault.Carron}@lip6.fr

**Abstract.** Specifications of Multi-Player Learning Games (MPLG) are expressed collaboratively by designers who don't speak the same conceptual language. Often, specifications contain design errors and inconsistencies that are difficult to detect in playtests. In this paper, we present a formal framework to assist designers in modeling and automatic verification of learning games at the design stage of development process.

**Keywords:** Multi-Player Learning Game, Learning Game Verification, Instructional Design, Game Design, Model Checking, Symmetric Petri nets.

## 1 Introduction

**Context.** Learning Games can be defined as "a virtual environment and a gaming experience in which the contents that we want to teach can be naturally embedded with some contextual relevance in terms of the game-playing [...]" [1]. Multi-Player Learning Games (MPLG) are learning games involving multiple players who are competitors or collaborators.

MPLGs are software applications resulting from costly and complex engineering processes, involving multiple stakeholders (domain experts, game designers, learning designers, programmers, testers, etc.) who often do not speak the same language. Usually, at the end of the development process of MPLGs, testing activities are conducted by humans testers who explore the possible executions of the game to detect both design errors and programming bugs. However, the design errors are particularly hard to find and so specifications and design properties become hard to verify. This observation is more true for MPLGs implying multiple players which are concurrent and dynamic systems implementing complex interactions between game universe and players.

**Problem.** First, the difference between the language used by game designers and the language of learning designers can lead to inconsistency and design errors in specifications. In this context, the construction of a rigorous specification clarifies several aspects of system behaviour that are not obvious in an informal specification. Then, the complexity and dynamic nature of MPLGs makes more difficult the verification of properties on specifications, only by playtests. Indeed, it is widely acknowledged that it

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 476–481, 2014.
© Springer International Publishing Switzerland 2014

is cost-effective to spend more efforts at the specification stage, otherwise, many flaws would go unnoticed only to be detected at the later stages of software development that would lead to iterative changes to occur in the development life cycle [2]. In addition, the exploration of all possible execution paths of MPLG scenarios is impossible manually due to their huge number. For example, it is difficult for learning designers to ensure, in game specifications, properties such as the winner of the game is always a player that has acquired all domain skills or to ensure that it is proposed systematically a reinforcement activity to a player who fails in a game level.

**Contribution.** We advocate to check the properties of MPLGs (such as those mentioned above) at the design stage before starting the programming stage in order to reduce cost of testing activities and verify these properties automatically. We propose a framework in order to formalize and verify game scenarios, at the design stage. Our objective is to ensure that MPLG satisfies properties which are extremely difficult to assess only by means of playtests. Thus, once the verification is performed on an abstract design, development starts from a validated design. After this formal verification, the test of MPLGs would be less costly.

This paper presents a formal approach enabling automatic verification of MPLG properties. Among the available techniques, we chose the *Petri nets* to formally specify MPLG and *model checking* techniques to verify properties.

Petri nets are a mathematical notation suitable for the modeling of concurrent and dynamic systems [3]. Due to the dynamic nature of learning games, we selected a particular Petri net model: Symmetric net with bags. Model checking is a powerful way to verify systems; it provides complete proof of correctness automatically, or explains, via a counter-example, why a system is not correct [4]. This counter-example can be used to pinpoint the source of the error [5].

**Content.** The following section details our verification framework. Section 3 presents conclusion and perspectives of our approach.

## 2 Verification Framework

Today, video game companies use human testers to detect errors in games. Obviously, this method is costly and unreliable (most of games receive several patches after their release date). In our approach, we assume that MPLG scenarios become more reliable and the development less costly if game specifications are verified prior to programming stage.

Among the multiple models adapted for modeling behavior of MPLG, we chose the Colored nets that are necessary to get a reasonable sized specification, thanks to the use of colors to model data. Next, within the large variety of colored Petri Nets, we selected Symmetric Nets with Bags (SNB) [6] where tokens can hold bags of colors. They support optimized model checking techniques [7] allowing to verify properties of MPLG. Moreover, the notion of bags is relevant to model some dynamic aspects (game objects can appear or disappear, knowledge of players can increase or decrease) that are typical of MPLGs in a much simpler way than with most other colored Petri nets.

We provide here a short and informal presentation of Symmetric Nets with Bags.

### 2.1 Symmetric Petri Nets with Bags (SNB)

### 2.2 Informal Definition and Example

A Petri net is a bipartite graph composed of places (circles), that represent resources (*e.g*, the current state of player in the game) and transitions (rectangles) that represent actions and consume resources to produce new ones. Some guards ([conditions] written near a rectangle) can be added to transitions.

The SNB of Figure 1 models activities of generic MPLG. Place `beforeActivity` holds players and their context: skills and virtual state (stored in bags). The initial marking `M` in place `beforeActivity` contains one token per player (identified by `p`) representing his skills and virtual state (sets `S` and `V` respectively). Place `activityDesc` holds the required skills and virtual states for each activity. The initial marking `M'` in place `activityDesc` contains a token per activity (identified by `a`) with its prerequisite (`S-In` and `V-In`) and the information needed to compute the consequence of the activity on the player (in terms of `S-Out` and `V-Out`).

Each activity begins (firing of transition `start`) only when player's skills and his virtual state include the prerequisite of the activity. Then, the activity may end in failure (transition `looseA`) or successfully (transition `winA`). Functions `fwin` and `floose` represent the evolution of skills and virtual states of players at the end of the activity (dropped in place `beforeActivity`). A player wins when `winCond` is true, it expresses conditions on skills or/and virtual state.

The SNB shown in figure 1 allows us to model with very abstract and concise manner MPLG scenario. This powerful expressiveness allows us to have the whole scenario on a "small" graph (useful for automatic execution) but for a better understanding, it is possible to imagine it "deployed": one SNB for each activity.

SNB preserve the use of symmetry-based techniques allowing efficient state space analysis [7] that is of particular interest for the formal analysis of MPLGs. The model in Figure 1 is exactly the one that is verified (once max values defined), it is not mandatory to instantiate it per activity and player. We have thus a powerful formalism and some tools [7] to verify the expected properties of a MPLG.

We have applied our automatic verification approach to multi-players, parallel, concurrent or collaborative scenarios of MPLGs in order to test the feasibility of our formal framework. We present the results obtained when studying the MPLG "Wonderland" .

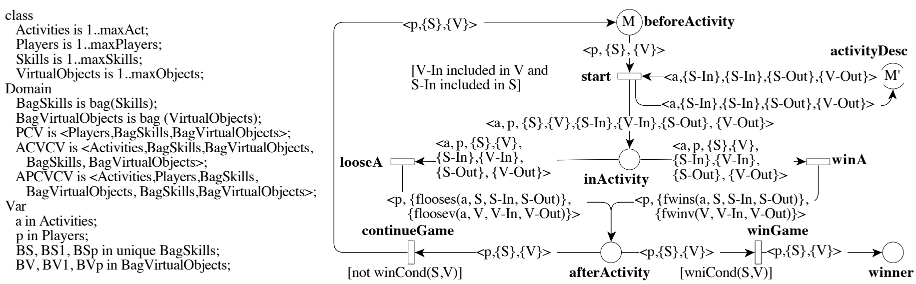



**Fig. 1.** Modeling a game activity in SNB

“Wonderland” aims to train pupils of the primary school to solve arithmetic problems. Players have to find and liberate a princess kidnapped by a dragon. To succeed, players have to resolve arithmetic problems. In term of skills, a player wins when he holds all the skills, therefore `winCond` is `card(S) = maxSkills`.

### 2.3 Verification

Once properties are specified for MPLG scenario, we formalize them for the automatic verification. In our case, we use (1) the temporal logic to formalize the properties and (2) both the CPN-AMI [8] and the Crocodile tool [7] as model checkers, able to process efficiently Symmetric Nets. These model checkers terminate (if we have sufficient memory resources and consider finite systems) with a positive answer or a negative one. If one property is not verified (negative answer), the model checker provides a counter-example. This counter-example is useful to pinpoint the source of the error and to correct the specifications. Since the formal representation of MPLG allows the construction of the reachability/quotient graph, we can verify properties automatically.

The properties presented in the table 1 are informally specified but they were described by temporal logic formulas [4] and model-checked automatically with Crocodile tool. It provided us either a complete proof of correctness of properties or a counter-example.

*Ending and winning properties.* Let us first define the property $winner(p)$ stating that a player won the game. The learning-dependent associated property provided by the learning designer is *“The player who acquires all skills wins the game"* (cf. the table 1). Therefore, $winner(p)$ is true if: $\exists\langle p,\{S\},\{V\}\rangle \in beforeActivity$ such that $card(S) = maxSkills$. In other words, there is a player token in place `beforeActivity` such that the set of the player’s skills contains all the possible skills.

We call $WinningProperty$ the property that ensures that a player can win the game: $WinningProperty \Leftrightarrow \exists p \in \texttt{Players}, winner(p)$ (corresponding to the property *“A player who acquires all skills wins the game"*). We call $EndingProperty$ the property identifying the end of the game. The game designer specified it as *“Only one player can win the game and all others lose it”* (cf. the table 1), it was formalized as follows: $\exists! p \in \texttt{Players}, winner(p)$.

The Crocodile tool has proved the first property and gave us counter-examples for the second one. It allowed us to detect an inconsistency in specifications between the game designers and the learning designers. Once the inconsistency detected, the designers must agree on the solution to implement. If they want to model a game where only one player can win (*i.e.* once a player wins, the others cannot begin a new activity) they must change the guard of transition `start`. This transition can be fired only if the marking of place `beforeActivity` does not contain a token $\langle p,\{S\},\{V\}\rangle$ such that $winner(p)$.

The learning-dependent property *“No player may liberate princess without all the skills"* and the game-dependent one *“The player wins only when he liberates the princess"* associate the winning condition to the liberation of the princess. Crocodile tool allows us to detect inconsistencies between the two winning views (acquisition of all the skills and liberation of the princess) and the specification of the game : a player could liberate the princess without acquiring all skills or he/she could acquire all the skills without liberating the princess. The model checker allowed us to detect these inconsistencies and to correct them in the learning game specifications.

**Table 1.** Wonderland properties

| Learning-dependent properties | Game-dependent properties |
|---|---|
| *"The learner can always improve his skills or at least maintain them", "One skill is seen by the learner not more than 2 times ", "The player who acquires all skills wins the game", "There is no a winner without all skills", "The player can not acquire the skill "complicated multiplication" as long as he does not have the skill "simple multiplication"* | *"The Hermit's telescope is used by at most one player", "Only one player can win the game and all others lose", The player wins only when he liberates the princess", "there is no deadlock between players", "The player has to meet the wizard before winning the telescope ladders", "The player has to help the hermit before winning the telescope", "The player has to kill the dragon before delivering the Princess"* |

*Scheduling properties.* Among the temporal properties we have verified, the Crocodile tools allowed us to detect a problem in specifications about the property *"The player can not acquire the skill "complicated multiplication" as long as he does not have the skill "simple multiplication"*. However, Wonderland SNB allows a sequence where the skill "complicated multiplication" is used whitout acquiring the skill "simple multiplication". Thanks to the model checker we detected this design error.

*Learning process property.* It aims to verify that a player always has the possibility to increase his skills (until he wins the game). Such a property is a temporal logic property since it is necessary to compare the states along each execution. We call *increaseSkillsProperty* the associated property.

We define first the $increaseStrictly(s,s',p)$ property where $s$ and $s'$ are two states and $p$ a player. $increaseStrictly(s,s',p)$ is true if the set of skills of player $p$ at state $s$ is strictly included in the set of skills of player $p$ at state $s'$.

Then, $increaseSkillsProperty = \forall p \in \texttt{Players}, \forall s$, a reachable state, the set of skills is equal to `Skills` or there is a path leading to $s'$ such that $increaseStrictly(s,s',p)$.

This important learning-dependent property would have been very difficult to be verified manually. The use of Crocodile tool allowed us to validate it in the Wonderland scenario.

## 3 Conclusion

We presented verification framework allowing the formal modeling and the automatic verification of MPLG scenarios. Our objective is to reduce cost of MPLGs development by enabling detection of specification errors and inconsistencies at the design stage. One interesting point of our approach is to provide tools helping both learning designers and game designers to verify the MPLG specifications. The originality of our work can be summarized in three points:

– our work aims to detect inconsistencies between learning designers and game designers who do not speak the same conceptual language. Designers may have difficulties to describe consistent and non-ambiguous specifications of MPLG scenarios;

- the formal framework that we propose addresses the learning and the game properties at the same time;
- the model of Petri Nets that we chose and its optimized model checking techniques allow us to verify specifications automatically in a finite time.

We applied our approach on specifications of MPLG "Wonderland" and we were able to detect inconsistencies and errors in properties specified by both learning and game designers.

**Future Work.** Other research are conducted to extend the framework in order to support the verification of properties related to collaboration and to concurrency in MPLGs. We are also working on an editor to assist designers in building SNB models. We have considered MPLGs composed of independent activities, we now have to put our framework to the test of other types of MPLGs and adapt it if necessary.

**Acknowledgements.** This work was supported in part by the Region Ile de France and by the french Ministry for the Economy, Industry and Employment (FUI). We would like to thank them for their support in the "PlaySerious" project.

## References

1. Fabricatore, C.: Learning and Videogames: An Unexploited Synergy. In: 2000 AECT National Convention. Springer Science + Business Media, Secaucus (2000)
2. Rajib Mall. Fundamentals of Software Engineering. Prentice-Hall of India Private Limited. Rajkamal Electric of India Private Limited
3. Jensen, K., Kristensen, L.: Coloured Petri Nets: Modelling and Validation of Concurrent Systems. Springer (2009) ISBN: 978-3-642-00283-0
4. Bérard, B., Bidoit, M., Finkel, A., Laroussinie, F., Petit, A., Petrucci, L., Schnoebelen, P.: Systems and Software Verification. Model-Checking Techniques and Tools. Springer (2001)
5. Clarke, E.M., Grumberg, O., Peled, D.: Model checking. MIT Press (1999)
6. Haddad, S., Kordon, F., Petrucci, L., Pradat-Peyre, J.-F., Trèves, N.: Efficient State-Based Analysis by Introducing Bags in Petri Net Color Domains. In: 28th American Control Conference (ACC), pp. 5018–5025. Omnipress IEEE (2009)
7. Colange, M., Baarir, S., Kordon, F., Thierry-Mieg, Y.: Crocodile: A Symbolic/Symbolic tool for the analysis of Symmetric Nets with Bag. In: Kristensen, L.M., Petrucci, L. (eds.) PETRI NETS 2011. LNCS, vol. 6709, pp. 338–347. Springer, Heidelberg (2011)
8. Hamez, A., Hillah, L., Kordon, F., Linard, A., Paviot-Adet, E., Renault, X., Thierry-Mieg, Y.: New features in CPN-AMI 3: Focusing on the analysis of complex distributed systems. In: 6th International Conference on Application of Concurrency to System Design (ACSD), Turku, Finland, pp. 273–275. IEEE Computer Society (June 2006)

# Guess It! Using Gamificated Apps to Support Students Foreign Language Learning by Organic Community-Driven Peer-Assessment

Anke Berns[1], Manuel Palomo-Duarte[2],
Juan Manuel Dodero[2], and Alberto Cejas[2]

1 Facultad de Filosofía y Letras
2 Escuela Superior de Ingeniería
University of Cadiz, Spain
{anke.berns,manuel.palomo,juanma.dodero}@uca.es,
alberto.cejassanchez@alum.uca.es

**Abstract.** Guess it! is a gamificated app for smartphones that aims to enhance students foreign language learning outside the classroom. Learning contents in the APP are not delivered to our learners but, rather than this, they are produced by versatile tasks (e.g. rating or reporting, adding new definitions, etc.) that although individually performed, affect the community of learners. Firstly, we detail how the app can be used to support students language learning outside the classroom, secondly, how it helps learners to get actively involved in their own learning process, thirdly, how the server intelligence is organically built on top of community contributions and peer-assessments, so the more students use the app, the richer the knowledge base becomes. And finally we show how the data on students usage of the app and participation in the learning process can be used to automate the assessment by the language instructor.

**Keywords:** mobile learning, foreign languages, gamification, apps, peer-assessment, demonstration, open-source software.

## 1 Pedagogical Background

The purpose of the current work is to provide students with those learning tools which are able to support and engage them towards foreign language learning, outside the classroom. For this purpose we have designed a gamificated app, called Guess it! which is based on the pedagogical framework of mobile learning. As outlined by several researchers [3] mobile learning theories recognizes not only the essential role of mobility and communication in the process of learning but also the importance of the context to construct meaning. Mobile learning has also been defined as a process of gaining knowledge through multiple contexts amongst people and personal interactive technologies. In line with the aforementioned aspects Sharples proposes a framework, based on Engeström's expansive

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 482–485, 2014.
© Springer International Publishing Switzerland 2014

activity model [1], to analyze the interdependencies between learning and technology. The framework he proposes is mainly based on five factors (subject, object, context, tools and communication) which are themselves analysed under two different perspectives: a technological and a semiotic perspective. The technological perspective describes learning as an engagement with technology, in which devices such as mobile phones function as interactive agents which helps its users to gain knowledge, to communicate amongst each others, to share and negotiate contents and meanings and finally, to reflect on them. In contrast, the semiotic perspective describes learning as a cognitive system in which learners objective-oriented actions are mediated by cultural tools and signs. Guess it! has been designed on the aforementioned framework and aims to offer learners with a highly interactive and flexible learning environment that allows them not only to ubiquitously get and produce new language input and output, but also to personalize their learning contents according to their needs and background knowledge. This is done by providing learners with several tasks that focus on students' reading, listening and writing skills. The tasks consist in guessing and explaining different terms with the help of definitions in the target language, but which are not necessarily related solely to the target language culture. The initial corpus of definitions is gradually created, assessed and reported by the players themselves, being formative [5,2,4]. In this way our learners are encouraged to participate actively in their own learning process reviewing and reflecting on the target language as well as the different contexts in which this is used. Hence learners are organically co-constructing the system knowledge. Furthermore, in order to make the different learning tasks more challenging, we gamificated the app adding features such as a score system and rankings.

## 2 Learning Design and Architecture

Guess It! follows a client-server architecture as we can see in figure 1. Client software is installed in students smartphones or tablets. It identifies the student and interacts with the server through the Internet.

The typical application workflow begins by asking the user to select a language and entering his login credentials to be redirected to the main menu. Once a player has logged in the game he can choose the level, categories and number of definitions he wants to play. In this way, the learning process is highly customizable according to the educational needs. Once within the game screen, the app will provide a definition of a term that the user will try to guess with the optional help of a clue. After entering the right answer or wasting 3 attempts, the system will show the term defined and the user will get the chance of listening its pronunciation and of adding the definition to his notebook for a later review. Then, the user has to rate the quality of the definition before passing to the next one. At the end of the game, an overall summary is shown along with some stats to gamificate the app. However, the "Statistics" option from the main menu, will preserve the last game data along with a bulk of additional useful information regarding all the games and definitions played.

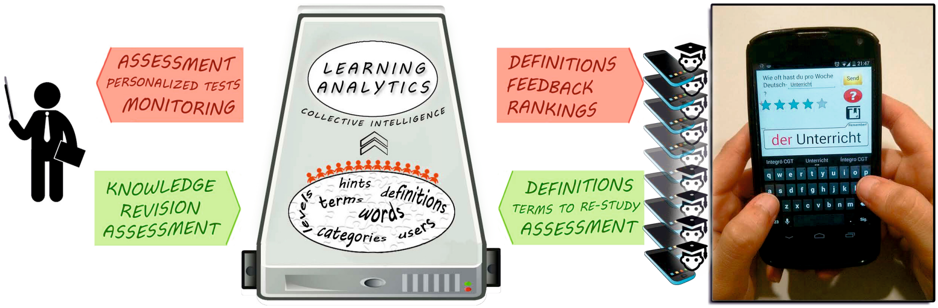



**Fig. 1.** Guess it! architecture

Additionally, for every 20 definitions played, the user is allowed to provide a new definition for a term. That term was previously introduced in the system by the teacher according to his learning objectives. This way the app does not only oblige the students to guess terms, but also to define them. Once introduced a new definition in the system, it will be played by the other students, whereas the ratings obtained for the definition become peer-assessments for the author.

The server program stores the knowledge base, which includes the following data: terms to study including their category and level, the definitions available for each term, the definitions each student did guess or fail during each game-session, the assessment of each player and the terms that each player considered interesting to be restudied. With that information the system provides its users with some useful feedback and different rankings, showing their learning progress as well as level compared with the ones achieved by other peers. Additionally this knowledge is being increased and continuously reviewed by the users themselves. As a result of this process, the server will contain different graded definitions for each term.

In order to monitor and assess the above mentioned process the course coordinator can have access to different learning analytics which include: the usage which has been made by each student of the app, the grades he obtained (according to the terms that has been guessed and the grades his definitions have received from peers), the grade he provided for each definition (compared with those of the rest of the players), a report of the low-graded definitions can provide insights of the difficulties of a certain student, and another report with those poor definitions that unfairly received high grades can highlight those students who are unable to detect mistakes in the target language.

Finally, the server can create personalized tests for each player. These tests might include those definitions that a student previously guessed and might even detect if he guessed them by cheating using a dictionary or if he guessed them by really having the required knowledge. Additionally, the personalized tests might also include a certain number of definitions that a student either did not guess or did not focus on during the different game sessions. Furthermore the system will provide the supervisor with different recommendations. For example, if the ratio of students guessing a certain definition is below a threshold, we have to check if the linguistic level of the definition is more complex than it should be or if the students perhaps need a reinforcement in some language topics.

## 3 Conclusions

For the current project we have specifically developed an Android APP (available in Google Play[1]) and a PHP server program that are available as GPL free software in its GitHub forge. From here users can download the system, see screenshots as well as screencasts and get sample sets of definitions in different languages and levels to start using the system in their own servers. The app has been designed to support students ubiquitous language learning and incorporates German, English and Russian language definitions. Nonetheless it can easily be used for any language and at any level by simply adding the suitable vocabulary and definitions in the knowledge base. Furthermore it takes advantage of the crowdsourcing possibilities of the player communities to improve students experience thanks to the collective intelligence implemented in the server. Our future work consists in implementing the system for a large number of students and to observe how it evolves and what is its impact on students' learning.

**Acknowledgments** This work has been funded by the Andalusian Government under the University of Cadiz programme for Researching and Innovation in Education and by the European Union under the OpenDiscoverySpace (CIP-ICT-PSP-2011-5) project. We would also like to thank the OSLUCA as well as Elihú Salcedo, Pilar Romero, Ignacio Calleja, Andrea Calderón and Sebastián Pitula for their support.

## References

1. Engeström, Y.: Learning by expanding: an activity-theoretical approach to developmental research. Orienta-Konsultit Oy, Helsinki (1987)
2. Gikandi, J., Morrow, D., Davis, N.: Online formative assessment in higher education: A review of the literature. Computers & Education 57(4), 2333–2351 (2011)
3. Sharples, M., Taylor, J., Vavoula, G.: Towards a theory of mobile learning. In: Proceedings of MLearn 2005 Conference (2005)
4. Taras, M.: Using Assessment for Learning and Learning from Assessment. Assessment & Evaluation in Higher Education 27(6), 501–510 (2002)
5. van Zundert, M., Sluijsmans, D., van Merrinboer, J.: Effective peer assessment processes: Research findings and future directions. Learning and Instruction 20(4), 270–279 (2010)

[1] https://play.google.com/store/apps/details?id=es.uca.tabu

# Towards Academic Analytics by Means of Motion Charts

Jan Géryk[1] and Lubomír Popelínský[2]

[1] Computer Systems Unit, Masaryk University, Brno, Czech Republic
[2] KD Lab Faculty of Informatics, Masaryk University, Brno, Czech Republic
{geryk,popel}@fi.muni.cz

**Abstract.** Academic Analytics facilitate creation of actionable intelligence to enhance learning and student success. Visual Analytics methods like Motion Charts show changes through time by presenting animations within two-dimensional space and by changing element appearances. We present enhanced user interface features of the Visual Analytics tool EDAIME primarily designed for exploratory analysis of Academic Analytics. The tool supports various visualization methods and implements Motion Charts adjusted to academic settings. We also describe rich user interface abilities in detail.

**Keywords:** Academic Analytics, Animations, Motion Charts, Student retention, Visual Analytics.

## 1 Introduction

Higher education institutions have a major interest in improving the quality and the effectiveness of the education. In [1], authors concluded that the increasing accountability requirements of educational institutions represent a key to unlocking the potential for Academic Analytics (AA) to effectively enhance student retention and graduation level. In [2], authors utilized analytics for developing predictive models of student enrollment and retention, and for identifying at risk students. They also highlighted three critical success factors–executives committed to evidence-based decision-making, staff with adequate data analysis skills and flexible and effective technology platform. However, authors warned that more elaborated accountability can raise issues with privacy, faculty executive's involvement, and data administration.

Data visualization is an important part of successful understanding of complex data. Visualization of multivariate data without fundamental dimensionality reduction remains a challenge. Animations represent a promising approach to facilitate better perception of changing values. In [3], authors pointed out that animations help to keep the viewer's attention. Motion Charts (MC) represents an animated data presentation method that shows multiple elements and dimensions on a two dimensional plane, as described in [4].

MC elements are basically a two-dimensional shape representing one object from the data set. The variable mapping is one of the most important parts of the exploratory data analysis and no optimal method for mapping the data variables to element positions and appearances is available. Naturally, some mappings are more

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 486–489, 2014.
© Springer International Publishing Switzerland 2014

appropriate and transparent than others. Both the data characteristics and the investigative hypothesis should influence the selection of a variable mapping.

MC allows exploring long-term trends which represent the subject of high-level analysis as well as the elements that form the patterns which represent targeted analysis. The dynamic nature of the MC allows better identification of trends in the longitudinal multivariate data and enables visualization of more element characteristics simultaneously, as presented in [5]. Several rules important for better clarity of transitions are presented in [6]. Some of them are also relevant for the efficient design of MC and the beneficial user interface features.

In this paper, we show importance of data visualization for successful understanding of complex data and we also characterize the MC concept. In the next section, we present several papers that successfully utilized MC. Subsequently, we describe rich abilities of the EDAIME user interface. Finally, we summarize the results and present future work.

## 2 Related Work

Number of papers concerning MC has increased recently. In [4], authors illustrated how MC can tell dynamic stories. They utilized data about current employment statistics and presented differences between the perception of common static charts and tables, and the dynamic manner of MC. Authors concluded that static presentation style serves well the purpose of relaying accurate quantitative data to the analyst. Authors also emphasized that the benefit of MC lies in displaying rich multidimensional data through time on a single plane with dynamic and interactive features. They also pointed out that users are allowed to easily explore, interpret, and analyze information behind the data.

Beneficial feature for better visual perception of changes in time-series analysis is presented in [7]. Author emphasized both the benefits and the drawbacks of common data visualization methods, namely line charts and bar charts. Then, the author focused on dealing with issues concerning the time-series analysis. Subsequently, the author concluded that patterns of change through time can take many meaningful forms and introduced new feature, called visual trails, designed for MC. The feature allows tracing the full path that values take from one point in time to another. The feature is used for overcoming visual perception limitations of motions and allows analyst to examine degree of change, shape, velocity, and direction of change. Finally, the author successfully demonstrated proposed improvements.

In [5], authors have presented SOCR Motion Charts infrastructure designed for exploratory analysis of multivariate data. The interactive data visualization tool enables the presentation of the high dimensional data. The tool is designed using object-oriented programming, implemented as a Java applet and is available publicly. Authors successfully validate the visualization paradigm using several publicly available data sets including housing prices or consumer price index.

## 3 EDAIME User Interface Features

We elaborately described the design and the implementation of the EDAIME tool in [8]. The EDAIME offers a vast selection of features suitable for analyzing of AA which are not available in common charts. The configurable interactive features facilitate analyst to explore the data in more. User interface features are customizable and allow analyst to fine tune the display and variable mapping according to his or her needs. Analysts can also select input data, arrange the information, and interpret the statistics behind the data. Available features include mouse-over data display, color representation, plot size representation, plot trails, animated time plot, variable speed of animation, changing of axis series, changing of axis scaling, distortion, and support of statistical methods.

Regardless of the great power of human brain, memory is quite limited. It is difficult to reconstruct the past events from memory, to recapture the sequence of events and details of each moment. When the trail feature is activated, the corresponding element is highlighted and the visual trail retains the entire path of the element as the time move forward and presents the passage of time as a single image. The trail is preserved during the whole MC animation and facilitates the easy visual tracking of the history of the element and enables to overcome the limitations of the memory. As described in [7], visual trails are particularly useful to reveal the nature of change and can help to examine the magnitude of change, shape of change, velocity of change, and direction of change. The nature of change of the corresponding trail is also calculated on the background and the trail is analyzed by statistical functions to provide additional information. The information reflects the statistical relationship between variables.

Support of statistical methods is particularly useful for examining the nature of change. The methods provide simple summaries that form the basis of the initial description of the data and also serve as a part of a more extensive analysis. We utilized both univariate and bivariate statistical methods. Method input parameters consist of MC variables. When the animation playback is running, all of the statistical measures are computed for every element. Mouse-click event on each element will extract the interactive HTML table on the right side of the chart. The table is consisted of computed measures for every element sorted in descending order according to the value.

Layout of the EDAIME user interface consists of menu C*ontrols* that enables controlling animation playback. Apart from play, pause, and stop buttons, there is also range input field which controls five levels of the animation speed. The controls facilitate the step-by-step exploration of the animation and allow functionality for transparent exploration of the data over the entire time span. Analyst can interactively march through time by traversing mouse over semester number localized in right bottom of the screen of the EDAIME tool. Mouse-over element events trigger tooltip with additional element-specific information. Menu *Data mapping* allows mapping data into MC variables. Univariate and bivariate statistical functions can be applied on any of the MC variables. List of univariate measures include coefficient of variation, skewness, mean, variance, standard deviation, median absolute deviation, median, geometric mean, and interquartile range. Bivariate measures include sample covariance, sample correlation, and paired t-test.

## 4 Conclusion and Future Work

Common data visualization methods have limitations in terms of the volume and the complexity of the processed data. Animations are substantially transparent procedures that can readily present a good overview of the complex data. MC presents multiple elements and dimensions of the data on a single two-dimensional plane. The main contribution lies in enabling critical questions about data relationships and characteristics.

In the EDAIME tool, we enhanced the MC concept and expanded it to be more suitable for AA. We also developed the rich user interface with impressive selection of features which provide a quick, convenient, and visually appealing way to identify potential correlations between different variables. The complexity of underlying data to be analyzed is reduced by the analyst's ability to control how the data is mapped to MC variables and displayed in the EDAIME tool. Among other, analysts can interact with the charts by refining chart criteria and clarify analyses of the relationship between variables.

The additional representation of the data provides the analyst more possibilities in exploring the data, but the additional functionality can be also confusing. To verify user friendliness and usability of the tool, we will carry out a controlled experiment with two groups of users. They will utilize different VA tool and methods trying to understand the same dataset.

## References

1. Oblinger, D., Campbell, J.P.: Academic analytics. Educause (2007)
2. Campbell, J.P., DeBlois, P.B., Oblinger, D.G.: Academic analytics: A new tool for a new era. Educause Review 42(4), 40–57 (2007)
3. Tversky, B., Morrison, J.B., Betrancourt, M.: Animation: Can It Facilitate? International Journal Human-Computer Studies, 247–262 (2002)
4. Battista, V., Cheng, E.: Motion Charts: Telling Stories with Statistics. In: JSM Proceedings, Statistical Computing Section, pp. 4473–4483. American Statistical Association, Alexandria (2011)
5. Al-Aziz, J., Christou, N., Dinov, I.D.: SOCR Motion Charts: An Efficient, Open-source. Interactive and Dynamic Applet for Visualizing Longitudinal Multivariate Data Journal of Statistics Education 18(3) (2010)
6. Heer, J., Robertson, G.: Animated Transitions in Statistical Data Graphics. IEEE Transactions on Visualization and Computer Graphics, 1240–1247 (2007)
7. Few, S.: Visualizing Change: An Innovation in Time-Series Analysis. In: Visual Business Intelligence Newsletter, White paper SAS (2007)
8. Géryk, J., Popelínský, L.: Visual Analytics for Increasing Efficiency of Higher Education Institutions. In: Proceedings of the 6th Workshop on Applications of Knowledge-Based Technologies in Business (accepted 2014)

# ILDE: Community Environment for Conceptualizing, Authoring and Deploying Learning Activities

Davinia Hernández-Leo[1], Juan I. Asensio-Pérez[2], Michael Derntl[3], Luis P. Prieto[4], and Jonathan Chacón[1]

[1] ICT Department, Universitat Pompeu Fabra, Barcelona, Spain
{davinia.hernandez,jonathan.chacon}@upf.edu
[2] School of Telecommunication Engineering, Universidad de Valladolid, Spain
juaase@tel.uva.es
[3] Advanced Community Information Systems (ACIS), RWTH Aachen University, Germany
derntl@dbis.rwth-aachen.de
[4] École Polytechnique Federale de Lausanne (EPFL), Switzerland
luis.prieto@epfl.ch

**Abstract.** This demonstration paper presents the Integrated Learning Design Environment (ILDE). ILDE is being developed in the METIS project, which aims at promoting the adoption of learning design by providing integrated support to teachers throughout the whole design and implementation process (or lifecycle). ILDE integrates existing free- and open-source tools that include: co-design support for teacher communities; learning design editors following different authoring and pedagogical approaches; interface for deployment of designs on mainstream Virtual Learning Environments (VLEs). The integration is designed so that teachers experience a continuous flow while completing the tasks involved in the learning design lifecycle, even when the tasks are supported by different tools. ILDE uses the LdShake platform to provide social networking features and to manage the integrated access to designs and tooling including conceptualization tools (OULDI templates), editors (WebCollage, OpenGLM), and deployment into VLEs (e.g., Moodle) via GLUE!-PS.

**Keywords:** learning design, integrated environment, teacher communities, authoring, deploying.

## 1 Introduction: Meeting Teachers' Co-design Needs

The METIS project ("Meeting teachers' co-design needs by integrated learning environments"; www.metis-project.org) aims at providing a complete set of conceptual and technological tools capable of supporting teachers along the whole process of going from pedagogical intentions to the setting up of all the technological resources required for their enactment. METIS builds on the body of research on learning design [1,2] that has been accumulated during the last decade, as well as on the myriad of available free and open learning design authoring tools that cover a broad range of design practices and pedagogical approaches. The ultimate METIS

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 490–493, 2014.
© Springer International Publishing Switzerland 2014

goal is to foster the adoption of learning design practices and tools in real educational practice, without disrupting the current technology-enhanced learning ecosystem, mainly dominated by Virtual Learning Environments (VLEs) and Web 2.0 tools. Thus, METIS's main intended outcomes are: an Integrated Learning Design Environment (ILDE) that enables teachers to choose among multiple learning design authoring tools, (co-)produce, explore, share, evaluate and implement [3] learning designs in mainstream VLEs (e.g., Moodle); and, a set of professional development (PD) actions (e.g., workshops) customized for different educational levels. Such PD actions are aimed at training educators in the use of learning design practices supported by the ILDE, and are also proposed as an instrument to eventually become part of the training programmes of educational institutions. METIS emphasizes the role of community/social aspects as change drivers in teaching practice [4].

This paper describes the current ILDE version (section 2), which has already been employed for supporting several PD actions on learning design (section 3).

## 2 ILDE: Integrated Learning Design Environment

The ILDE supports cooperation within "learning design" communities in which their members share and co-create multiple types of learning design solutions covering the complete lifecycle. This has been achieved by the integration of a number of existing learning design tools (Fig. 1), supporting:

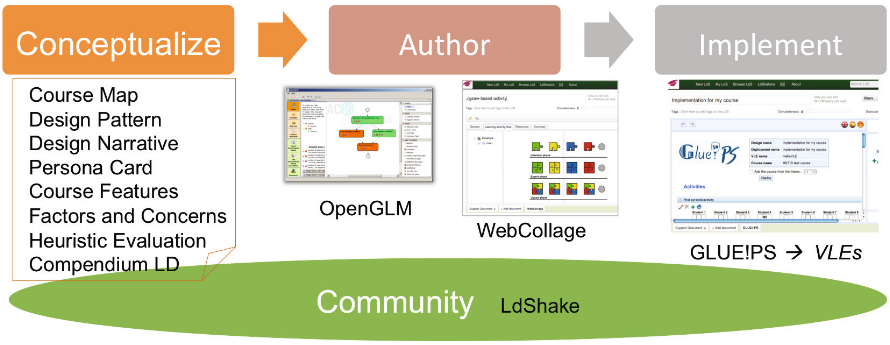



**Fig. 1.** Schema of tools integrated in the ILDE

**Community:** To support co-design, sharing, self- and peer-review evaluation of learning design solutions the ILDE is built on top of LdShake [4]. LdShake provides social network features, including sharing designs with different access rights, public comments to designs, private messages, exploration of shared designs, exploration of community members' activity, etc. Thus, it also acts as a repository, controls the access to designs, and embeds in its web space conceptualization, authoring and implementation tools; all in all, offering an integrated user experience.

**Conceptualization:** Before starting with the actual creation of learning designs, it is important to reflect about the characteristics of the context in which the designs will be applied (e.g., Personas, Factors and Concerns), sketch ideas for the design (e.g., Course Features, Course Map) and reflect about abstract descriptions (e.g., Design Pattern, Design Narrative). The ILDE integrates a number of design templates and tools facilitating the conceptualization of learning design solutions, most of them derived from the OULDI project [5].

**Authoring:** Authoring is the key step between the often sketchy conceptualization of a learning design solution and the implementation of the learning design solution in VLEs that provide the runtime environment for the design artefacts. This requires the production of a detailed definition of a learning design that can be deployed and executed with a specific group of learners. There are several tools available for this learning design stage. Most of the tools allow visual authoring of learning designs, yet each tool provides its own authoring approach. For instance, WebCollage [6] allows learning design modelling in the web browser based on design patterns; being a web-based application, WebCollage is directly integrated within ILDE's workflow. OpenGLM [7], on the other hand, is a desktop application that can produce learning designs in a format understood by the ILDE, which can be used for implementation of such detailed designs. To integrate with the ILDE, OpenGLM offers search in, import from, and export to ILDE within its desktop user interface. Similarly, CADMOS [8], being currently integrated in ILDE, separates the structural-conceptual perspective on learning design (e.g. activities) and the flow perspective. Generally, ILDE allows the attachment of arbitrary file types to the design solutions, some of which can be used further for implementation – e.g., IMS Learning Design [9] compliant packages.

**Implementation:** This phase includes a first step in which an authored learning design is particularized for a concrete learning situation, e.g., a “course” in a specific VLE. It involves creating groups with the students enrolled in the VLE, assigning groups to different learning activities, and selecting learning tools to support those activities (both tools provided by the VLE itself, or external third-party web 2.0 tools that are integrated in the VLE using the GLUE! system [10]). For carrying out this particularization of the designs, the ILDE makes use of the GLUE!-PS system [11]. GLUE!-PS translates learning designs, represented with the computational languages of the different authoring tools, into a common internal representation or *“lingua franca”*. The teachers, using the GLUE!-PS graphical user interface, then manipulate these representations. In a second step, once all the details of the particularized design are worked out, GLUE!-PS “deploys” it into the VLE, i.e. it sets up and configures all the VLE elements that represent the learning design (activities, groups, tools, …). Thus, for instance, a WebCollage authored learning design can be automatically transformed into a Moodle course ready to be accessed by the participating students.

## 3 Pilots Using the ILDE

The ILDE has been employed in the support of several PD actions in institutions for Vocational Training, Adult Education and Higher Education. Part of the PD actions consisted of “collaborative learning (CL) training workshops” in which participants were instructed about CL fundamentals, practised in co-conceptualizing and co-authoring CL activities, and were finally guided through the process of deploying their designs into Moodle. All workshop activities were carried out with the support of the ILDE, including the sharing and peer-evaluation of the designs. A second subset of PD actions included the enactment of a selection of the designs with students in authentic conditions. These enactments close the cycle of going from pedagogical ideas to their actual use with students, with the support of the ILDE. A total of 48 teachers were involved in a first round of workshops (12/2013-4/2014). A second round of workshops is planned for the second half of 2014.

**Demonstration Outline.** The main ILDE features and affordances (as described in section 3) will be demonstrated at the conference. A video of the ILDE and links to the different installations used in the workshops (mentioned in section 4) are available at http://ilde.upf.edu/about/.

**Acknowledgments.** This work has been partially funded by EACEA, METIS Project 531262-LLP-2012-ES-KA3-KA3MP and the Spanish EEE project (TIN2011-28308-C03-03). The authors acknowledge the contributions from other project members, including the technical assistants: Pablo Abenia, Javier Hoyos and Rizwan Uppal.

## References

1. Conole, G.: Designing for Learning in an Open World. Springer, Heidelberg (2012)
2. Mor, Y., Craft, B., Hernández-Leo, D.: The Art and Science of Learning Design. Research in Learning Technology Supplement 21, 22513 (2013)
3. Hernández-Leo, D., Chacón, J., Prieto, L.P., Asensio-Pérez, J.I., Derntl, M.: Towards an Integrated Learning Design Environment. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds.) EC-TEL 2013. LNCS, vol. 8095, pp. 448–453. Springer, Heidelberg (2013)
4. 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)
5. 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/
6. 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 67, 79–97 (2013)
7. 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)
8. Katsamani, M., Retalis, S.: Designing a Moodle course with the CADMOS learning design tool. Educational Media International 49, 317–331 (2012)
9. IMS Global: Learning Design (2003), http://imsglobal.org/learningdesign
10. Alario-Hoyos, C., Bote-Lorenzo, M.L., Gómez-Sánchez, E., Asensio-Pérez, J.I., Vega-Gorgojo, G., Ruiz-Calleja, A.: GLUE!: an architecture for the integration of external tools in virtual learning environments. Computers & Education 60(1), 122–137 (2013)
11. Prieto, L.P., Asensio-Pérez, J.I., Muñoz-Cristóbal, J.A., Dimitriadis, Y., Jorrín-Abellán, I.M., Gómez-Sánchez, E.: Enabling Teachers to Deploy CSCL Designs across Distributed Learning Environments. IEEE Transactions on Learning Technologies 6(4), 324–336 (2013)

# Med-Assess System for Evaluating and Enhancing Nursing Job Knowledge and Performance

Marjan Khobreh[1], Fazel Ansari[1], Mareike Dornhöfer[1], Réka Vas[2], and Madjid Fathi[1]

[1] Institute of Knowledge Based Systems,
Department of Electrical Engineering & Computer Science,
University of Siegen, Germany
{marjan.khobreh,fazel.ansari,m.dornhoefer,
madjid.fathi}@uni-siegen.de
[2] Corvinno Technology Transfer Center, Budapest, Hungary
rvas@corvinno.com

**Abstract.** The European funded project Med-Assess supports assessing of work-based competences and job knowledge of nurses, indicating existing knowledge gaps, and ultimately providing recommendations for improving nursing competences. This paper presents the Med-Assess concept, and reflects the implementation results of its ontological approach for analysis and assessment of nursing job knowledge. The ontological approach matches the nursing requirements and domain specific knowledge, and provides the logic for assessment of the end-users i.e. job applicants, nurses and care-givers.

**Keywords:** Knowledge Base, Assessment, Ontology, Nursing, Med-Assess.

## 1 Introduction

Health sectors across Europe demand highly qualified nurses. Nursing education is strongly tailored to on-the-job training and learning by doing. The key to determine quality of nursing performance is determined as the ability and skills for applying domain specific knowledge in various situations properly. Work-based competences can be categorized in three groups: competences as (i) a list of tasks, (ii) a collection of attributes, and (iii) a holistic or integrated relationship [1]. In a wide overview meta-competence is concerned in relation with two key points, firstly, “how to apply skills and knowledge in various task situations”, and secondly, “how to acquire missing competences”. Assessment of the competence level is associated with job performance. The assessment is to assure attainment of the work objectives, or to measure capabilities for reaching higher performance level. Competence assessment is classified into two groups, (i) assessment of objective competence with standard test of job knowledge requirements, or (ii) subjective competence (task-based competence) with performance assessment. Assessing competences of nurses is at the core of the EU-funded Innovation Transfer project Med-Assess: Adaptive Medical Profession Assessor [2]. Med-Assess aims at emphasizing “competences required” and “competences taught/learned” in order to analyze “competence gaps”, and consciously recommend

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 494–497, 2014.
© Springer International Publishing Switzerland 2014

"learning materials"[3]. In order to apply competence assessment and competence gap analysis, a meta-ontology is formalized for Med-Assess to clarify what is the relation between (a) nursing tasks demanded by health sectors, (b) competences required for performing certain tasks, and (c) learning needs, which enable the practitioners to perform the task [3]. Med-Assess can be utilized within the job to assess the level of competence regularly, analyze the competence gaps, and improve their knowledge to fill these determined gaps. A nurse acts the role of learner, and at the same time she performs the given tasks as an expert (supervisor) for novice/beginner nurses. This is a learning process for developing the nurse's personal profile through learning on the job. Alike, the job applicants and nursing students can use Med-Assess to estimate their competences based on the market demands.

## 2 Med-Assess System

Med-Assess deploys a knowledge-based approach [3] with the main feature being the Med-Assess Knowledge Base (MAKB). It uses different types of ontologies to establish the semantics between certain entities. Ontology represents "domain-specific entities along with their definitions, a set of relationships among them, properties of each relationship, and, in some cases, a set of explicit axioms defined for those relations and entities" [4]. STUDIO, an in-house software solution, provides features for creating ontologies and storing the test questions and learning materials [5]. The ontology model is the backbone of the STUDIO system that, on one hand, provides a formal description of the domain of interest, on the other it serves as a basis of the adaptive knowledge testing solution. The Nursing Ontology –stored in the Ontology Repository of STUDIO – is the central element of the content authoring processes in Med-Assess. The structure of content is also determined by the ontology in STUDIO, meaning that every piece of content (learning material or test questions) is connected to one specific concept of the Nursing Ontology. Learning materials are stored in the *content repository*, while test questions are stored in the *test bank*. Additionally, there is also a *content developer* application, which offers content management functions to the content authors; and a *content presentation* component, which is entitled to present the stored content pieces (both tests and learning materials) to the users.

The already mentioned knowledge base (MAKB) consists of (a) **Medical domain** provides the formalization of relations between concepts/terms of the medical knowledge domain, which a nurse should know, the (b) **Task domain** provides the formalization of concepts/ terms of the nursing tasks in form of taxonomy and the (c) **Meta model** provides the relation between the Medical and Task domains. The main modules of Med-Assess are (1) ***U****ser profiling,* which manages and stores all information about a user, which is provided by the system or entered by the user, (2) ***A****ssessment* of user competences, analyzing the competence gaps and (3) ***R****ecommendation* for recommending appropriate learning materials based on detected gaps. The *User's Profile Data Base* (*UPDB*) is another feature and establishes a repository to store the user's information, test results or competence level. The first step in this module is to collect the information of a user as curriculum vitae (CV) of the applicant and consequently storing it in the *UPDB*. After taking the test and providing the recommended learning materials in the *A* and *R* module, the second step are the outputs, as the test result and the list of recommended learning materials will be delivered not only to the user, but also stored in *UPDB*. *Assessment* and *Recommendation* aim at assessing the applicant's competence level and recommending the appropriate learning materials for improving competences gaps. *A* and *R* consist of *MAKB*, *Test Engine* and *Recommender*. The main requirement for the building of the test question structure is that the questions are about the knowledge of a nurse, and it should be possible to apply an algorithm for calculating a score and automatically recommending matching learning content based on the candidates test results. Med-Assess currently works with multiple choice questions, where the candidate has to answer which of the named statements or described processes is correct or wrong, which statement does fit the given context the most or least or how the nurse would act, given a certain situation. Of course the questions and respected answers are specified by domain experts.

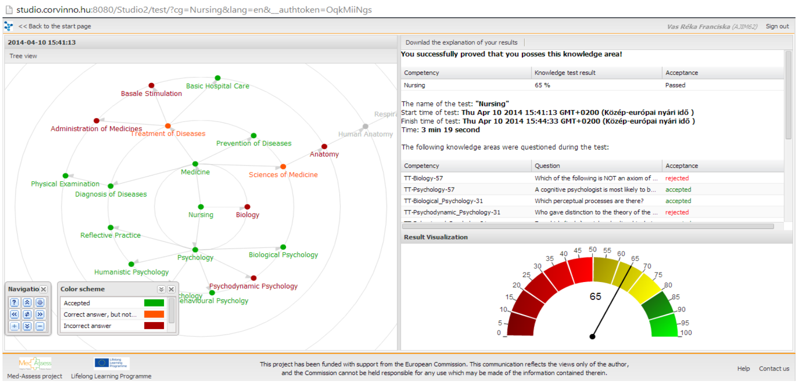



**Fig. 1.** Screenshot of Med-Assess Competence Assessment Dashboard

Taking into account the question styles, a short explanation of the *Adaptive Test Engine* and *Evaluation* module follows. The *Adaptive Test Engine* generates test packages depending on the competence of the applicant. As mentioned earlier, Med-Assess is built from a nursing ontology, where each node has got a number of test questions and related learning content. If a candidate takes the test, she starts at a more general level (of the ontology), and is led to the detailed level based on her answers. This is built like a ring structure (c.f. Fig. 1, left side). The inner circles are asking for the more general knowledge like nursing in general, while the more difficult or expert questions are on the outer rings. A candidate is only able to access these outer questions, when she is able to answer the more general ones before. The answers are categorized in three levels of accepted, not totally accepted and rejected. The *Test Engine* calculates the test score in a numerical way: first the percentage of correct answers is calculated on the outer most rings and these percentages are summed up in the next (inner) level. Summing is continued until the most inner knowledge element is not reached. In case that the answer is not correct or not totally accepted (c.f. Fig 1), the system provides a detailed explanation why the candidate failed. The explanation is concerning that knowledge element (i.e. which question or questions were incorrectly answered), and a link of related learning material describing that specific knowledge element is also provided.

## 3 Conclusion and Future Work

The present paper introduces the features and the background for developing the Med-Assess system. Med-Assess raises practical implications for the health sector in Germany. For the improvement of the system, the major issue is to develop and standardize a reference process model of nursing for adopting Med-Assess with various protocols and curricula in Germany as well as autonomies of nursing across EU. In addition, nurses may develop their career profile by practicing the lessons in various situations. Therefore former experiences of nurses are a valuable source for enriching the MAKB. The experiences should be extracted e.g. using structured interviews, classification and validation by experts. This issue has also been discussed in the former publication (c.f. [6]). The performance of the nurses should also be analyzed by empirical analysis of results through participating in several test sessions. Last but not least, technological, pharmacological and process-related changes may oblige a nurse to improve her knowledge, learn new terminologies, and practice new tasks. The problem of new-knowledge for new-tasks, and evidences of changes need to be carefully taken into account for sustaining the Med-Assess system with regard to advances in treatment and changing market demands. These are considered as open issues and potential works for further development of Med-Assess in future.

**Acknowledgement.** The research presented in this paper has been funded with support from the European Commission under the grant no. DE/12/LLP-LdV/TOI/147557. This communication reflects the views only of the authors.

## References

1. Guerrero, D., De los Ríos, I.: Professional competences: a classification of international models. J. Procedia - Social and Behavioral Sciences 46, 1290–1296 (2012)
2. Med-Assess – Adaptive Medical Profession Assessor, Official homepage of Med-Assess Project, http://www.med-assess.eu
3. Khobreh, M., Ansari, F., Dornhöfer, M., Fathi, M.: An ontology-based Recommender System to Support Nursing Education and Training. In: German Conference on Learning, Knowledge, Adaptation (LWA-2013), Bamberg, Germany (2013)
4. Shah, N., Muse, M.: Ontologies for Formal Representation of Biological Systems. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies - International Handbooks on Information Systems, Springer, Heidelberg (2009)
5. STUDIO – Ontology Driven Learning Environment, Official homepage of STUDIO, http://corvinno.com/web.nsf/do?open&lang=en&page=proj-studio#
6. Khobreh, M., Nasiri, S., Fathi, M.: E-Nursing: Experience Platform for Improving Nursing Performance. J. Geoinformatics 10(1), 57–63 (2014)

# Integrated Environment for Learning Programming

Martin Labaj, Marián Šimko, Jozef Tvarožek, and Mária Bieliková

Slovak University of Technology in Bratislava,
Faculty of Informatics and Information Technologies, Ilkovičova 2, 842 16 Bratislava, Slovakia
{martin.labaj,marian.simko,jozef.tvarozek,
maria.bielikova}@stuba.sk

**Abstract.** In this demo paper we present adaptive educational system ALEF, which addresses several drawbacks of existing systems supporting learning programming such as limited support for specific adaptation, collaboration, and motivation resulting from complexity of learning programming, which must involve practicing and active experimentation. ALEF constitutes comprehensive annotation framework to support interaction and collaboration, motivation component for keeping learners engaged and extensible architecture allowing to include additional programming languages. ALEF has been used for 5 years at Slovak University of Technology and has served more than 1 200 students.

**Keywords:** learning programming, integrated learning environment, ALEF.

## 1 Challenges in Support of Learning Programming

There are many domains where adaptive educational systems are successfully employed to support learning (many of them presented at conferences such as EC-TEL, ITS/AIED, UMAP). The advances in technology, web technology in particular, in the recent decade brought substantial advantages for interaction and collaboration [3]. It enabled interactive learning environments to emerge in order to support domains where learning from 'static' materials does not suffice. An example of such domain is learning programming. According to Kolb, a complete learning cycle includes active experimentation [6]. This particularly applies for learning programming, where theoretical mastering of (often abstract) programming concepts is not enough.

Many of educational systems created to support programming focus on automatic assessment of programming tasks, social support or adaptation during learning. Automatic assessment systems are widely used in university programming courses. Beyond managing submissions of programming assignments, returning feedback on correctness and assigning grades, these tools have evolved from command-line interfaces to comprehensive web-based tools that begin to integrate with learning management systems [5]. Recently, web-based programming support systems emerged to better support learning programming at various levels (writing code, debugging, analysing) by utilizing Web 2.0 features (e.g., [4]). Notable advances in adaptive support for learning programming can be found in works of Brusilovsky [2].

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 498–501, 2014.
© Springer International Publishing Switzerland 2014

However, despite recent advances in technology enhanced learning programming, learners in majority of programming supporting educational systems may still encounter several drawbacks:

- *Limited support for adaptation.* Often no proper user model is built, user actions are stored on per-document (not per-concept or per-competence) basis.
- *Learning object type monotony.* Often only one learning object type is present on screen at a time. Explanations or exercises are presented separately.
- *Limited reusability.* Existing systems provide access to isolated materials in terms of reusability and shareability.
- *Lack of collaboration and active learning support.* No or just limited means to communicate, share ideas and collaborate on problem solving are often present. Learners are rarely involved in content creation and enrichment.
- *Poor motivation.* Despite the recent advancements, support for learner's motivation is not properly developed and/or integrated to the whole learning cycle.

In this demo paper we present educational system ALEF [8], which addresses aforementioned drawbacks. ALEF is a web-based adaptive learning environment with a strong emphasis on collaboration and learner's active participation in learning process. It encourages learners to be not only passive consumers of information, but also to communicate with peers, collaborate on problem solving and contribute to learning content and/or metadata [1]. Even though it was created as a general framework for support of adaptive learning, most courses provided during 5 years of its operation are within the domain of learning programming, which requires specific support.

## 2 ALEF: Integrated Adaptive Learning Environment

The content in ALEF is presented via learning objects (LOs): (i) explanations (explaining topics), (ii) questions (ranging from single-choice to plaintext input), and (iii) interactive exercises (a particularly important from the learning programming perspective). The user model (learner's knowledge model, based on lightweight domain model [9]) is inferred from the interaction with exercises, questions, reading explanations, using collaborative functions and other aspects of the system. It serves as a basis for *adaptation engine* employing different adaptation strategies [1].

**Annotation Framework**. The heart of ALEF's social side is its annotation framework consisting of in-text interaction and presentation, sidebar annotation presenter and annotation browsers (widgets) [1]. Widgets are displayed around and within the content as small windows suited for different purposes, e.g., allowing the learners to chat in groups, annotate the content via pop-up widget displayed after selecting a text, tagging the content, or navigating through it using the menu widget, adding links to external sources, or adding specific annotations such as own definitions of key concepts. Interactive content is also presented in widgets and as such, it can be included into other content, e.g., interactive exercises can be shown in explanations.

**Extensible Architecture.** ALEF's architecture was designed to allow interconnection with external systems and services to further improve learning experience. Connecting the learning programming environment with other environments requires multiple

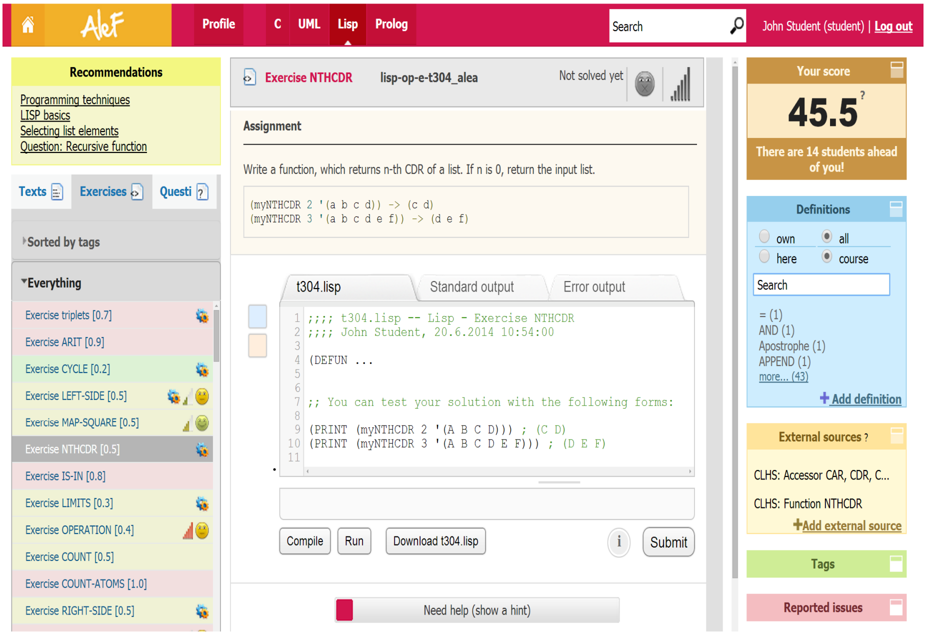


**Fig. 1.** Screenshot of ALEF. It is divided to three vertical parts: (i) navigational part (left); (ii) learning object with submission evaluator; (iii) learning and collaboration supporting widgets: score, term definitions, and external resources, other widgets (tags, issues in content) are shown collapsed

levels of cooperation. First, the user interface is shown in ALEF through the widgets (Fig. 1). A *proxy integration widget* simply points the learner to another learning environment, e.g., collaborative UML tool, while showing notifications about news, task assignments, peers, etc. from the other system. A *navigation integration widget* interfaces with the menu widget and allows selecting the content (e.g., another type of LO) presented within ALEF layout, but provided by different learning system. A *content integration widget* interfaces directly with the ALEF content, augmenting it with new features, e.g., adding a code editing and submission tool to ALEF exercises. External systems communicate relevant implicit and explicit user feedback to the ALEF for motivation rewards (user score, badges) and in order to augment the central user model, which is in turn provided to interconnected systems as a service.

We can observe an exemplary utilisation of ALEF's extensible architecture in incorporation of external services. ALEF has been integrated with external services at the university that provide interactive source code editing and submission evaluation [7]. It allows learners to write a program directly in the exercise widget and then compile it, test it using sample test inputs pre-defined in the exercise or using own test cases, and submit it for evaluation. In the case of incorrect submission, the learner is not only informed about any compile or run-time errors in the standard way, but also specific messages are configured for various passing/failing tests (e.g., cases gradually covering sample inputs from the exercise, then boundary cases, then large inputs) hinting the user what may be wrong with their code without revealing the test data.

**Motivation Component.** A successful execution of learning tasks in ALEF is rewarded with badges and adds points to the user score motivating the learner to progress through the course. Learner motivation is maintained also through dynamic score, which reflects activity of the whole class.

**ALEF in Practice.** ALEF supports education at our university since 2009 by providing adaptive interactive learning materials for 4 courses (Tab. 1). Students use ALEF both for homework and for practice during labs. They use annotation features for tagging and sharing text highlights or notes. We have an evidence of students' improvements both informally by monitoring their involvement and formally based on evaluation of tests results. We continually improve ALEF capabilities, mainly its collaborative and interactive features. E.g., we provided recently a component allowing learning from existing questions and answers provided by students and involving students into the answer correctness evaluation applying crowdsourcing principles.

**Table 1.** ALEF usage statistics (since 2009)

| Course - programming | Learning objects | Students participating | Learning object visited | Exercises solved | Questions answered | Annotations created |
|---|---|---|---|---|---|---|
| Functional | 308 | 146 | 29 121 | 2 106 | 3 878 | 3 843 |
| Logic | 235 | 127 | 12 004 | 988 | 370 | 2 061 |
| Procedural | 795 | 2 081 | 200 436 | 4 744 | 21 267 | 10 103 |
| UML | 694 | 1 184 | 434 163 | 2 290 | 12 366 | 40 377 |

**Acknowledgement.** This work was partially supported by the grants No. KEGA 009STU-4/2014, VG1/0675/11, VG1/0971/11 and APVV-0208-10.

## References

1. Bieliková, M., Šimko, M., Barla, M., Tvarožek, J., Labaj, M., Móro, R., Srba, I.: ALEF: from Application to Platform for Adaptive Collaborative Learning. In: Recomm. Systems for TEL, Springer Sci+Business Media, New York (2014)
2. Brusilovsky, P.: Adaptive Hypermedia for Education and Training. In: Adaptive Tech. for Training and Education, pp. 46–68. Cambridge University Press (2012)
3. Downes, S.: E-learning 2. Elearn magazine 10(1) (2005)
4. Hwang, W.-Y., Wang, C.-Y., Hwang, G.-J., Huang, Y.-M., Huang, S.: A web-based programming learning environment to support cognitive development. Interacting with Computers 20(6), 524–534 (2008)
5. Ihantola, P., Ahoniemi, T., Karavirta, V., Seppälä, O.: Review of recent systems for automatic assessment of programming assignments. In: Proc. of the 10th Int. Conf. on Comp. Educ. Res., pp. 86–93. ACM (2010)
6. Kolb, D.A., Boyatzis, R.E., Mainemelis, C.H.: Experiential learning theory: Previous research and new directions. Perspectives on Thinking, Learning, and Cognitive Styles 1, 227–247 (2001)
7. Návrat, P., Tvarožek, J.: Online Programming Exercises for Summative Assessment in University Courses. In: CompSysTech 2014. Springer, Ruse (2014) (accepted)
8. Šimko, M., Barla, M., Bieliková, M.: ALEF: A framework for adaptive web-based learning 2.0. In: Key Competencies in the Knowledge Society, pp. 367–378. Springer, Heidelberg (2010)
9. Šimko, M.: Automated Acquisition of Domain Model for Adaptive Collaborative Web-Based Learning. Inf. Sciences and Tech. Bulletin of the ACM Slovakia. 4(2), 1–9 (2012)

# ClipIt: Supporting Social Reflective Learning through Student-Made Educational Videos

Pablo Llinás[1], Pablo Haya[2], Miguel A. Gutierrez[1], Estefanía Martín[1], Jorge Castellanos[1], Isidoro Hernán[1], and Jaime Urquiza[1]

[1] Universidad Rey Juan Carlos, Madrid, Spain
[2] IIC, Universidad Autónoma de Madrid, Madrid, Spain
{pablo.llinas,miguelangel.gutierrez,estefania.martin, jorge.castellanos,isidoro.hernan,jaime.urquiza}@urjc.es, pablo.haya@iic.uam.es

**Abstract.** ClipIt is an online social learning platform focused on reflective learning using videos created by students. ClipIt does not follow a fixed methodology, nor is classroom based or teacher-led. Instead it responds to the online distance education paradigm, promoting both social and collaborative interaction and inducing to follow activity steps with subtle guidance which allows for maximum flexibility for the teacher. This paper presents the ClipIt platform and supporting a student-driven reflective learning process, specially focusing on social and collaborative interaction.

**Keywords:** CSCL, E-Learning, Personal Learning Environment, Reflective Learning, Service-Oriented, Social Collaboration, Video-based Learning.

## 1 Motivation

Web-based learning is emerging as the preferred e-learning option due to universal access, increased flexibility and multimedia capabilities [1]. However, facilitating access to learning resources is not enough to acquire a meaningful understanding. As it is well known learning performance depends on student's motivation. The JuxtaLearn project[1] provides a pedagogical and technological framework that encourages students to use creative activities, in particular video-making, to help them understand topics they find difficult to grasp. This approach provokes student reflection combining "learning by teaching" and ˝learning by doing".

In a typical e-learning environment, the two most extended roles are the learner and the author or teacher [2]. However, Personal Learning Environments (PLEs) [3] are changing these fixed roles allowing for consumers themselves to become producers, through creating and sharing. The learner has an active role based on a constructivism approach that takes place in social environments [4]. Furthermore, social interaction is also enhanced by non-task contexts [5], which affect the psychological

[1] http://juxtalearn.eu

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 502–505, 2014.
© Springer International Publishing Switzerland 2014

level of socialization, not so focused on task-achieving. In terms of facilitating social mechanisms, studies have stated an improvement in the quality of direct communication when using rich-media tools [6]. This is obtained by relating spontaneous or encouraged user discussion to learning media in a visual way, like YouTube promotes video-related discussion by locating comments right beneath them.

The e-learning scene is densely populated solutions (e.g. Moodle, Sakai, Blackboard, etc.) that begin to expose flaws with their all-in-one solution model [7]. Following the ideals of supporting collaborative, open and customizable platforms based on light-weight and extensible systems [8], we believe that e-learning platforms require to be service-oriented, instead of all-in-one solutions. This means that they should not be thought of as "where all the work is done", but as "one of many tools to help with my actual work". This is why "*future e-learning platforms must support the exchange of control between interoperating services*" [7].

We have sought to develop a web-space that manages data storage and interoperation between e-learning components. Students participate in a reflective learning process that is supported by an ecology of technical and pedagogical tools that combine social and collaborative learning principles. We devised a service-oriented architecture rather than following the "all-in-one" mainstream e-learning platform solution.

## 2 Social and Groupware Functionality

Combining the principles from PLEs, we devised a collaborative learning and social interaction system we named ClipIt. Social software and groupware tools are the two keystones of ClipIt, supporting networks of people, content and services [3][8]. It is based on the open-source social platform Elgg[2]. Its web front-end has been redesigned replacing the native Elgg web interface with our own in order to customize the platform's functionality and underlying methods and data structures.

ClipIt guides a learning process which revolves around a teaching concept, and which begins by planning out a storyboard to later create a video explaining this concept. Students work in groups for this goal, and ClipIt encourages group-work by featuring communication and sharing tools at a group-level. Each group have their "private" space in which to discuss about the composition process of storyboards and videos.

The resulting material is published into the activity scope, where all students working from any group in the activity are able to review and give feedback on their peers' material. In this stage, material-centred interactions occur between content producers and consumers. Teachers have an organizing and moderating role, but there is no linear work-path they must guide. This process of composing, publishing and getting feedback is repeated as necessary until each group has a final version of their video as a result of the activity. The requirements for making ClipIt a social-reflective platform can be divided into three categories: technical (TR), groupware (GR) and social (SR).

[2] http://elgg.org/

- **TR1 - User account control** for authentication and authoring of all data.
- **TR2 - Data interchange mechanisms** for storing and retrieving information and content produced/required by external tools using an API.
- **TR3 - Interoperability functionality** enabling to switch control over different services.
- **TR4 - Registering all events related to user interaction** as well as all inner operations which have an effect on content or user metadata.
- **GR1- Discussion and participatory production** enhancing facilities including file-exchange, contextualized communication, and peer-review evaluation tools.
- **GR2 - Moderating capabilities** for teachers upon the work carried out by students.
- **GR3 - Reflective feedback** encouraging metrics for self-awareness and positive feedback.
- **SR1 - Non-task context social interaction** promoting personal bonding and affiliation indirectly affecting the actual work by improving collaboration.
- **SR2 - Tagging mechanism** for linkage between the knowledge concepts and the produced material (storyboards, videos and any other file used in between).
- **SR3 - Sharing and recommendation mechanisms** to support dissemination of learner-created contents beyond the activity scope where the material was produced.

## 3 Use Case

Several trials with students have been carried out as part of the design and testing stages required to create ClipIt. In the first one, the main task was the creation of student-made videos related to “Java inheritance”. 40 students from the Object Oriented Programming subject (OOP) at the Universidad Rey Juan Carlos in Spain took part in this ClipIt activity. Students organized themselves in 17 groups, 8 of them composed by 3 students, 7 groups of 2, and 2 groups of 1. Two of the groups abandoned the trial, so in the end a total of 15 groups finished it.

All students were summoned to a briefing where they received instructions about the purpose of the trial, its length, and expected outcome. Teachers gave students some recommendations on how to make videos, performance-wise. They were given 14 days for delivering a video script, carry out the recording and editing process, and handing in the finished Video. When all the videos had been published, students had three days to give feedback by commenting and rating on the videos created by their peers. Once the activity was over, all of the students were summoned to a focus group where they carried out an experience feedback quiz regarding the ClipIt platform and the trial as a whole.

Students worked partly inside the ClipIt Web front-end, and partly using other methods of communication, but they ended up using ClipIt to store all of the intermediate material. In this trial, the learner had an active role based on a constructivism approach and was involved in an educational experience between peers. This benefits students as they are in a safe educational environment acquiring knowledge in a social and collaborative way. Fig. 1 is a snapshot of the dashboard for a user inside ClipIt. It gives an overview on the latest news, personalized for each student in a timeline with notifications about the running activities (left), pending tasks for the user and related groups (top), metrics of group progress and involvement in the activity (centre and bottom) and published videos (right), which represent the key learning material and are the final product created by students.

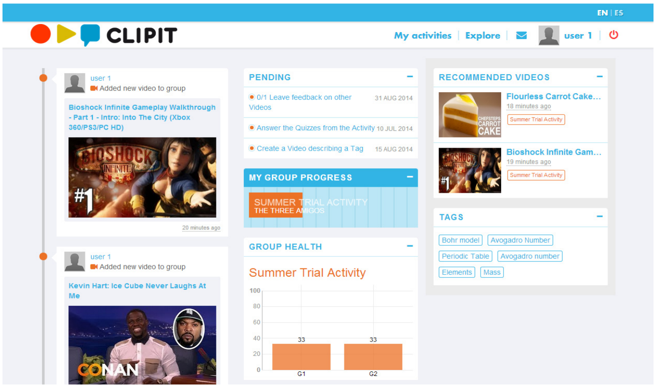



**Fig. 1.** ClipIt Web User Interface

**Acknowledgments.** The research leading to these results has received funding from the European Community's (FP7/2007-2013) under grant agreement n° 317964 JUXTALEARN.

## References

1. Horton, W.K.: Designing web-based training: How to teach anyone anything anywhere anytime. Wiley (2000)
2. Vossen, G., Westerkamp, P.: E-learning as a Web service. In: Proceedings. Seventh International Database Engineering and Applications Symposium, pp. 242–249. IEEE (2003)
3. Attwell, G.: Personal Learning Environments-the future of eLearning? ELearning Pap 2, 1–8 (2007)
4. Vygotsky, L.S.: Mind in society: The development of higher psychological processes. Harvard university press (1980)
5. Wegerif, R.: The social dimension of asynchronous learning networks. J. Asynchronous Learn. Netw. 2, 34–49 (1998)
6. Kreijns, K., Kirschner, P.A., Jochems, W.: Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. Comput. Hum. Behav. 19, 335–353 (2003)
7. Dagger, D., O'Connor, A., Lawless, S., Walsh, E., Wade, V.P.: Service-oriented e-learning platforms: From monolithic systems to flexible services. IEEE Internet Comput. 11, 28–35 (2007)
8. Van Harmelen, M.: Personal Learning Environments. ICALT, 815–816 (2006)

# PLATE Workbench: A PBL Authoring Tool

Yongwu Miao[1,*], Mohammed Samaka[2], and Disi Wang[1]

[1] Department of Computational and Cognitive Sciences, University of Duisburg-Essen, Duisburg, Germany
{miao,wang}@collide.info
[2] Computer Science and Engineering Department, Qatar University, Doha, Qatar
samaka.m@qu.edu.qa

**Abstract.** For most practitioners it is problematic to design an online or blended problem-based learning (PBL) unit within an increasingly complex learning context. This paper presents the development and application of a PBL authoring tool. It was based on a PBL-specific modeling language that offers expressive power through appropriate notations and abstractions focused on and restricted to PBL. In comparison with IMS-LD authoring tools, it provides more specific support for the design of online or blended PBL units.

**Keywords:** PBL, learning design, IMS-LD, PBL authoring tool, PBL script.

## 1 Introduction

Problem-Based Learning (PBL) is increasingly popular. By bringing real-life context and technology to the curriculum through a PBL approach, students are encouraged to become critical thinkers, problem solvers, team workers, and lifelong learners. As information and communication technologies (ICTs) advance, many ICT tools are used to support the implementation of PBL processes. However, for most practitioners it is actually problematic to design and conduct an online or blended PBL unit within an increasingly complex learning context.

Our general goal is to research and develop an innovative approach and an integrated PBL environment that support the design and delivery of online and blended PBL units with high flexibility and a low threshold of usage requirements. As a milestone towards this goal we have developed a web-based PBL authoring tool, called PBL Workbench. This tool is based on a PBL-specific modeling language that can be used to specify a wide range of PBL strategies as PBL scripts, computational descriptions of PBL processes. Then a PBL script can be transformed into a unit of learning (UoL) represented in IMS Learning Design (IMS-LD) [2], which can be executed in an extended IMS-LD run-time environment. In comparison with other IMS-LD authoring tools such as Re-Course [1] and Prolix OpenGLM [4], the PBL Workbench can provide more specific support for the design of online or blended PBL units.

* Corresponding author.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 506–509, 2014.
© Springer International Publishing Switzerland 2014

## 2 The Development and the Usage of a PBL Authoring Tool

We have developed the PBL Workbench, a web-based software application tool to support the design of various forms of online or hybrid PBL units. We adopted a domain-specific modeling language (DSML) approach to support PBL. In fact, applying DSML approach to develop an educational modeling language (EML) is not totally new in the field of technology-enhanced learning. Several EMLs have been developed in the past decade. However, rather than PBL, the target domains of these EMLs are more generic. Among them, IMS-LD [2], an e-learning technical standard, is a reference in the domain of education. It is important to note that IMS-LD is still very abstract and generic because its constructs are, for examples, “activity”, “property”, “data-type”, “expression”, “environment”, and “learning object”. Such a pedagogy-neutral modeling language cannot responsively support PBL. Because IMS-LD compatible authoring tools such as Re-Course [1] and Prolix OpenGLM [4] use such abstract concepts, it is not easy for most practitioners to develop, understand, and customize a PBL unit and to find appropriate mappings from PBL-specific concepts to generic and technical concepts. It requires sound pedagogy knowledge and overhead technical knowledge to represent complex teaching-learning processes [3] like PBL.

Through specializing the concepts of IMS-LD and introducing PBL-specific concepts with appropriate constraints, we developed a PBL scripting language on the top of IMS-LD. From the perspective of process modeling, we identified PBL constructs from many widely used PBL models and best practices through analyzing commonalities and differences of the semantics of the vocabularies used to describe various PBL strategies. Based on the rules to restrict and extend IMS-LD, a PBL script, a description of a PBL strategy represented in the PBL scripting language, can be transformed into an executable unit of learning (UoL) represented in IMS-LD. With the PBL authoring tool, a practitioner can design various forms of PBL, save them as PBL scripts, and then transform them into UoLs. Finally, a UoL transformed from a PBL script can be instantiated and executed by practitioners to scaffold PBL practices in an IMS-LD player such as SLeD [5].

The tool is implemented by using JavaScript both in client side and in server side (NodeJS). JSON is used to represent information as a unified data format in the system and MongoDB is used to manage and retrieve data. The tool is a web-based rich-client application. Fig. 1 shows a screenshot of the tool used to edit a PBL unit. The user interface of the tool consists of five parts. The menu bar on the top lists basic function and the state bar on the bottom indicates the current edit state. The central area contains the file manager (on the left), the graphic edit space (in the middle), and the property edit panel (on the right).

An online or hybrid PBL unit can be created by clicking the “*Course Design*”. After typing a title in a dialog window, the user can create a new PBL unit and then specify properties of the PBL unit such as the referred *organizational model, learning objectives, the driving problem, the description, learning outcomes, subject areas, grade,* and *recommended time frame*. A PBL script is consisted of a macro-script and several micro-scripts. A macro-script represents the high-level process that is made up of a set coordinated phases. The user can create and define a phase by dragging &

dropping a phase node and choosing one or more types from a list of pre-defined phase types such as *preparation, problem engagement, problem definition, idea generation, knowledge-gap identification, aim, plan, information sharing, investigation, reasoning, problem resolution, evaluation, application, reflection,* and *report*. To complete the definition of a phase, it is needed to specify the properties of the phase. The user can define the sequence of the phases by creating links between phases. In fact, it is allowed to define branch and loop structure through the use of the links as well.

Through clicking a phase node "*identify problem and set learning goal*" in an example PBL script, the user can edit and view the definition of the micro-script of the phase. As Fig. 1 shown, the user can create and specify a set of activities that can be performed in sequence, in branch, in parallel, or in loop. Two kinds of activity sequences can be specified: suggested sequence and controlled sequence. While a suggested sequence (represented by their presentation sequence, locations on the diagram) can be followed or not when conducting the PBL, a controlled sequence (represented with an explicit link between activities) will be controlled in a way only if the preceding activity completes, the succeeding activity starts. The user can create an activity by dragging and dropping an activity node and choosing an activity type from a list of activity types that are collected from the recommended activity types defined in the selected phase types. For example, in the phase *problem definition,* the authoring tool will recommend appropriate activity types such as *identify problem* and *formulate problem*. The activity can be further specified by assigning values of activity properties such as *learning objectives, description, coaching questions, work mode, collaboration mode, interaction mode, learning settings, used communication tools, start conditions, completion conditions,* and *recommended execution time.*

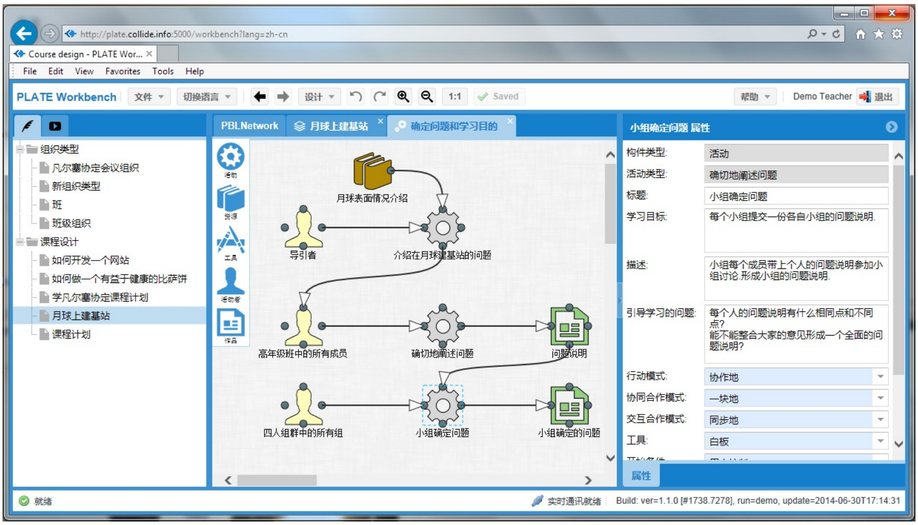



**Fig. 1.** An example of the micro-script of a PBL unit

The diagram shown in Fig. 1 represents a segment of the micro-script in the phase of "*identify problem and set learning goal*". First, the facilitator presents to the all students in a class a learning resource that introduces the Moon. Then all students identify and formulate problems individually. Then all students in groups discuss their ideas about problems. In the discussion the students in each group synthesize their ideas with whiteboards and document a problem statement. After that, they may submit problem statements to the facilitator or present in turn in the class. Fig. 1 shows the screenshot in which the activity "*formulate problem in groups*" was editing.

## 3 Conclusions and Future Work

In this paper, we briefly presented our approach to support for PBL. In particular, we presented a PBL authoring tool for the design of an online or blended PBL unit. By adopting a domain-specific modeling approach, the PBL authoring tool has been designed and implemented to support PBL through guiding the user to create appropriate types of phases, activities, artifacts, tools, resources and their relations. In comparison with IMS-LD authoring tools, it enables the design of online and blended PBL units with high flexibility and a low threshold of usage requirements.

Currently we are developing and extending the run-time environment. Because PBL is a self-regulated learning process, we will integrate the design-time function and run-time function to make it possible to dynamically change the PBL process model on the fly. It is also challenging to support similarity-based process information retrieval for reusing PBL scripts.

**Acknowledgement.** This paper was made possible by NPRP grant # NPRP 5 - 051 - 1 – 015 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

## References

1. Griffithsa, D., Beauvoira, P., Libera, O., Barrett-Baxendaleb, M.: From Reload to Re-Course: learning from IMS Learning Design implementations. Distance Education 30(2), 201–222 (2009)
2. IMS Learning Design, `http://www.imsglobal.org/learningdesign/`
3. Miao, Y., Koper, R.: An Efficient and Flexible Technical Approach to Develop and Deliver Online Peer Assessment. In: Proceedings of the 7th Computer Supported Collaborative Learning Conference, New Jersey, USA, July 16-21, pp. 502–511 (2007)
4. 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)
5. Weller, M., Little, A., McAndrew, P., Woods, W.: Learning Design, generic service descriptions and universal acid. Educational Technology & Society 9(1), 138–145 (2006)

# vCL: Platform for Customizing Individual and Group Learning

Yuliyan Nakov and Veronika Kozareva

ProZnanie JSC, 52 Gen. Gurko St., Sofia, Bulgaria
{ynakov,vkozareva}@proznanie.bg
http://proznanie.bg

**Abstract.** Technological innovations are gradually supporting the learning process and bringing new added values by changing pedagogical interaction. This paper aims to present vCL – a synchronous online interactive system, developed and tested specifically for the needs of distance learning. vCL provides easy access to a group of people as well as an opportunity for customizing individual and group learning. It is seen as an integration of technologies and pedagogy aimed at enhancing a collaborative learning effect. A special focus is put on vCL's social and psychological aspects in line with collaborative learning. The main limitations are listed and options for their mitigation are proposed. Finally, the basic conclusions of the prototype testing are systemized. During the demonstration, the conference participants will be able to take part in simulated sessions, and to see how vCL features could be applied in the learning process.

**Keywords:** synchronous virtual learning environment, distance education, collaborative learning, customized learning, interaction, videoconferencing.

## 1 Introduction

Traditional forms of education are increasingly integrating technological innovations, resulting in changed educational practices and methodology. These trends lead to the widespread use of distance learning solutions and their diversification. In recent years, synchronous (real-time) e-learning systems are proliferating in an effort to simulate the learner's conventional learning environment experience while building on the web-environment benefits.

There are a number of products on the market which make it possible to provide pedagogical interaction in real time. The majority of these systems, however, have to be redesigned and adapted to suit the needs of the education process rather than those of business, for which they were originally created.

This paper will present VCL – an online interactive system, developed and tested specifically for the needs of distance learning. It enables easy access to a large number of people (both participants and observers), as well as the customizing of individual and group learning.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 510–513, 2014.
© Springer International Publishing Switzerland 2014

## 2 Technological Innovation for Collaborative Learning

By definition, collaborative learning is a pedagogical method in which students work together to achieve a common goal, exchange views, clarify the meaning of concepts or solve problematic situations together. [1] Applied within a virtual environment, this approach is associated with an active process of collective construction of knowledge, using the group as a source of information, motivational agent and means of mutual support.

vCL is a browser-based platform for conducting training sessions in real time by live video-conference connection. The users (participants and observers) can access the system through HTML5-based application, distributed by PHP server. Supported by Javascript this application gives an opportunity for drawing, chat, documents processing, etc. The application connects with a Realtime Media Server for sending or receiving audio/video streams and mixes the individual video screens in a grid through a Media recorder. The communication between individual users, the media server and the data base is implemented through a control server, which sends the communication messages and records the sessions for future reproduction.

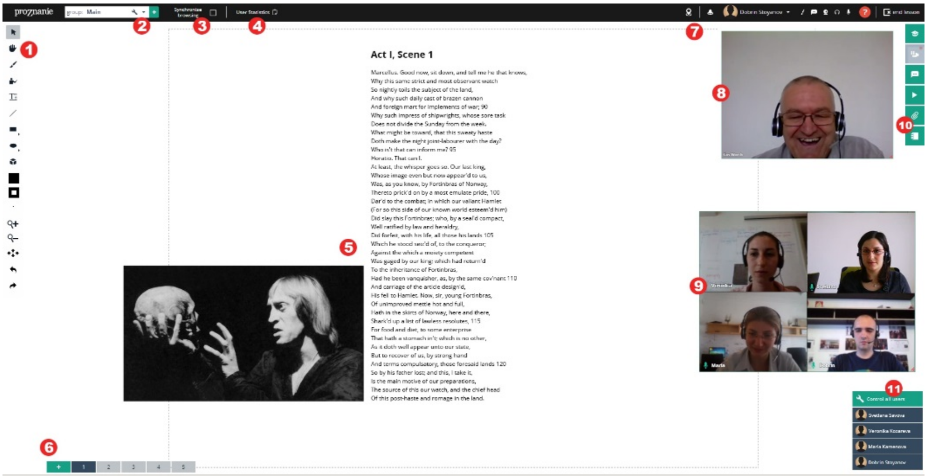



**Fig. 1.** – Example of presenter's screen view – 1. Whiteboard tools; 2. Groups' dropdown menu; 3. Synchronized mode; 4. User statistics; 5. Interactive whiteboard with pedagogical resources; 6. Work sheets; 7. Personal profile controls and information; 8. Presenter's video; 9. Participants' video grid; 10. Functional boxes menu (video screens, chat, media player, upload box, personal notes); 11. Individual and group controls

In addition to the video-conference connection, vCL offers a set of features for implementation of collaborative learning activities, such as an interactive whiteboard for content sharing, peer review, collaborative creation and editing of text, images, objects; media player for video and audio demonstrations; breakout rooms for individual work or work in small groups; teacher's control panel for managing group discussions, upload box for file sharing, etc. All these features bring a social aspect [2] to the virtual learning environment [3], making it interactive and favorable for conducting collaborative learning.

The flexible interface, the functionalities' management, as well as the different levels of access in the system (presenter, participant or observer), provide the opportunity to apply various pedagogical approaches. This makes vCL applicable in different forms of education/training: all stages of formal education, non-formal, informal or blended learning, as well as in a broad non-educational context.

## 3 Social and Psychological Aspects of the Online Interactive System

The design of the collaborative approach allows for an enhanced process based not only on the real-time communication, but also on the innovative use of specific system features as a new standard in learning experience. The vCL platform also allows for overcoming many of the limitations imposed by the traditional pedagogical reality – such as time, location, costs and a limit of the number of participants. For example, regular classes in English were held through the systems with adults participating from different locations. Moreover, the system opens access to distant virtual educators and representatives of different social groups and geographical locations. This is a prerequisite for the diversification of participants' groups and the representation of a wide range of individual perspectives in the collaborative learning process.

The psychological advantages of the system are connected mostly with overcoming stress and minimizing the time for social adaptation in the group. When the system was tested with groups of children, who did not know each other, a repeating trend was the rapid elimination of barriers in communication with both the teacher and the peers. A prerequisite for this is the comfort of the well-known (often home) environment during participation in the training and the absence of the usual physical distractions from the conventional classroom.

In a traditional learning environment, placing students in groups and giving them instructions does not mean collaboration will naturally happen [4]. In a virtual learning environment the coordination of group interaction, so that each participant contributes equally to the task performance, is essential. vCL provides the teacher with a better system for managing the individual and group work than in a real classroom.

Monitoring is another crucial factor for effective interaction and group dynamics management in the virtual learning environment [5]. vCL has a function for monitoring in real time, as well as for recording and archiving live sessions. These features allow analysis, ongoing diagnostics and external control of the learning sessions.

## 4 Limits

Essential for achieving meaningful group dynamics and interaction in the platform is an understanding not only of its advantages, but also of its disadvantages. This is also related to identification and involvement of strategies for overcoming the limitations.

On the one hand the computer-supported environment protects the private space, but on the other limits the overall performance of the personality. The options for adequate support in critical situations are also limited. This risk can be overcome with further research and improvement of social and pedagogical aspects of the system. Another approach to compensate the lack of physical contact with the teacher is to implement a methodologically rationalized system for blended learning.

Every participant must have at least basic computer skills in order to be able to participate fully. Therefore, we have tried to develop a completely intuitive interface that allows effortless operation and does not need any special explanations or training. In addition, our team has prepared self-learning tutorials to help the users get acquainted with the system and make most of it.

Reaching the limits of the system‘s functional abundance may create barriers to applying various pedagogical approaches. Our aim is to develop vCL features swiftly in the course on the platform’s operation.

## 5 Conclusions

In this paper we have presented vCL – a novel interactive educational platform for collaborative learning. The system’s prototype has been tested in various educational contexts since the beginning of 2014 – university courses, foreign language courses, science classes with children, company trainings, etc. The collected feedback information confirmed the effectiveness of the system regarding individualization of the learning process and achievement of a balanced group interaction. Currently, the platform is utilized and constantly customized according to the users’ ongoing feedback.

During the demonstration, the conference participants will be able to take part in simulated sessions with a distant virtual teacher and a group of students. The interaction will be structured in various collaborative activities which illustrate how vCL features could be applied in the learning process. A special focus will be put on the added value of technological innovation in the pedagogical interaction.

## References

1. Hron, A., Friedrich, A.: A review of web-based collaborative learning: Factors beyond technology. Journal of Computer Assisted Learning 19(1), 70–79 (2003)
2. Kreijns, K., Kirschner, P.A., Jochems, W., Van Buuren, H.: Measuring perceived sociability of computer-supported collaborative learning environments. Computers & Education 49, 176–192 (2007)
3. Martin, F., Parker, M.A., Deale, D.F.: Examining Interactivity in Synchronous Virtual Classrooms. The International Review of Research in Open and Distance Learning 13(3) (2012)
4. Johnson, D., Johnson, R.: Learning together and alone: Cooperative, competitive, and individualistic learning, 4th edn. Needham Heights, Allyn & Bacon (1994)
5. Wang, Q.: Design and evaluation of a collaborative learning environment. Computers & Education 53, 1138–1146 (2009)

# Illustrating a Graph Coloring Algorithm Based on the Principle of Inclusion and Exclusion Using GraphTea

M. Ali Rostami[1], H. Martin Bücker[1], and Azin Azadi[2]

[1] Institute for Computer Science, Friedrich Schiller University Jena, 07737 Jena
{a.rostami,martin.buecker}@uni-jena.de
[2] Jovoto Company, Berlin
aazadi@gmail.com

**Abstract.** Graph theory is becoming increasingly important in various scientific disciplines. The graph coloring problem is an abstraction of partitioning entities in the presence of conflicts. It is therefore among the most prominent problems in graph theory with a wide range of different application areas. The importance of graph coloring has led to the development of various algorithms to cope with the high computational complexity of this NP-complete problem. A comprehensive understanding of existing algorithms is not only crucial to future algorithm research in this area, but also beneficial to the education of students in mathematics and computer science. To better explain sophisticated graph coloring algorithms, this paper proposes the educational tool GraphTea. More precisely, we demonstrate the design of GraphTea to illustrate a recent graph coloring algorithm based on the principle of inclusion and exclusion. Our experiences indicate that GraphTea makes teaching this algorithm in classroom easier.

## 1 Introduction

The application of graph algorithms in various fields of science, industry, and society has been continuously increasing over the recent years and decades. Graph theory provides a powerful set of techniques to find an abstraction of a given problem. Furthermore, the visual representation in terms of graphs often gives an improved view of a problem leading to a more comprehensive understanding. GraphTea [5] is an extensible interactive software for teaching concepts from graph theory. In GraphTea, various algorithms for the solution of different graph theoretical problems are supported. Students can easily construct, modify, and run a rich set of different algorithms on instances of various graph classes.

Several educational tools with a focus on graphs have been developed in the past [2–4]. They address various concepts and include a diverse set of features. So, each tool has certain advantages and shortcomings. To the best of our knowledge, there is no other software than GraphTea with the aim to combine education, visualization, and rich-editing capabilities.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 514–517, 2014.
© Springer International Publishing Switzerland 2014

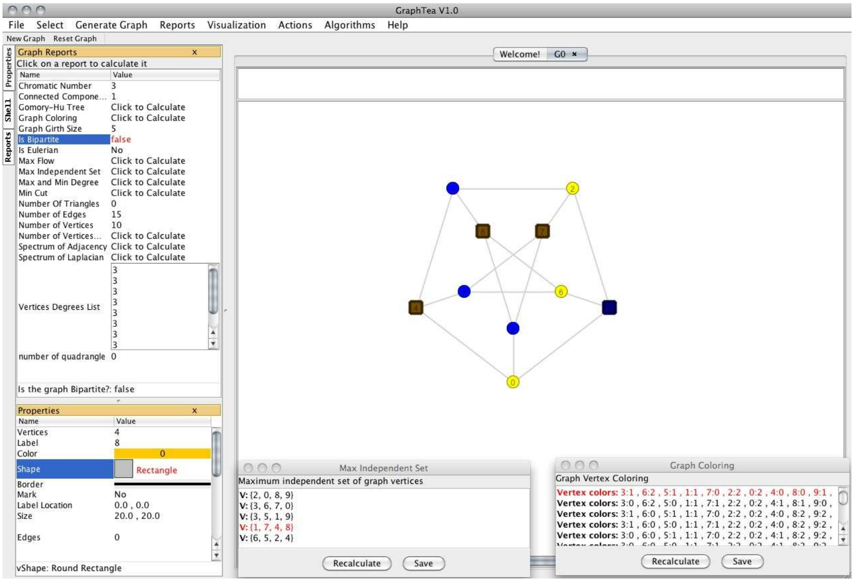



**Fig. 1.** An overview of GraphTea and two reports on a given graph

Throughout this demonstration paper, we focus on the graph coloring problem to illustrate teaching with GraphTea. Currently, the fastest exact algorithm for coloring a given graph is given in [1]. The idea behind this recent algorithm is based on the principle of inclusion and exclusion which leads to the following definition. Let $A$ and $B$ denote two sets whose elements are also sets. Their product, $A \times B$, is then defined as

$$A \times B := \bigcup_{a \in A, b \in B} a \cup b.$$

If $I$ is the set of all independent sets of a graph $G$, the coloring algorithm [1] consists of computing $k$ powers of $I$. The chromatic number of $G$ with a vertex set $V$ is then the smallest $k$ such that $V \in I^k$. The purpose of the present demonstration paper is to illustrate this algorithm using GraphTea.

## 2 The Graph Teaching Software GraphTea

GraphTea is a graph teaching software designed specifically to visualize and explore graph algorithms interactively. Figure 1 shows a snapshot of the main graph window together with two additional windows in the bottom right. These separate windows that provide additional information are called “reports.” Reports are explained in more detail in [5]. The two reports in Fig. 1 give the solutions to the maximum independent set and graph coloring problems.

An algorithm is a step-by-step procedure to achieve a goal. In classroom, a teacher tries to clarify each step of an algorithm, providing students proper and accurate illustrations as well as explanations that help to comprehend the algorithm. To mimic such an educational process, we added to GraphTea the novel feature of attaching an instruction to a step of the algorithm. In fact, such an instruction can be thought of as an "active comment" on a specific part of the algorithm. An instruction serves as an interactive clarification of a procedure. For instance, an instruction can be used to select a vertex, an edge, or a subset thereof. Another example is the option to enter an input that can be used to control the algorithm.

## 3 Interactively Exploring the Graph Coloring Algorithm

The algorithm to color a graph $G$ computes the powers of the set of all independent sets of $G$. Recall that $I$ denotes this set of all independent sets of $G$. Hereafter, we represent the product $I \times I$ by a list of rows as follows. Each element $c \in I \times I$ is represented by a separate row that lists the two non-empty distinct sets $a, b \in I$ where $c = a \cup b$. That is, the product $I \times I$ is visualized by a list of rows of the form $a$, $b$ where $a$ and $b$ are both elements of $I$; see Fig. 2 (Middle) for an example. However, since $a$ and $b$ should be colored differently, we omit in this list those rows whose two sets have a non-empty intersection.

GraphTea provides a novel way to explore and visualize the whole algorithm. After the graph generation, the coloring algorithm can be executed by selecting the name "inclusion-exclusion coloring algorithm." The so-called algorithm runner window shows the generation of the set of all independent sets.

When the student pushes the button "Play One Step" in the top of the algorithm runner window, the set $I \times I$ is computed and visualized. The rows in this list are selectable and a selected element from $I^2$ is highlighted in red in this list. The sets of vertices in the graph that correspond to this selection are colored in the graph. Each set of vertices is assigned a unique color.

This procedure of computing powers of $I$ is repeated until the set of vertices $V$ is an element of $I^k$ for some $k$. This is indicated by at least one row in the list representing $I^k$ that contains all vertices from $V$. Any row that consists of set of vertices whose union equals $V$ describes a coloring with a minimal number of colors. For this sample graph, the algorithm terminates with $k = 4$. This final situation is illustrated in Fig. 2 (Left). As soon as the final results are obtained, all rows in the list of $I^k$ that represent a minimal coloring become bold as shown in Fig. 2 (Middle). The algorithm runner window is used throughout the execution of the algorithm to monitor the progress of the algorithm. In Fig. 2 (Right), the status of the algorithm runner window is illustrated upon the termination of the algorithm. This window shows the separate steps of the algorithm beginning with the initial computation of $I$. Then, the algorithm runner window prints a line whenever another power of $I$ is computed. Finally, it also reports the chromatic number which is equal to 4 in this example.

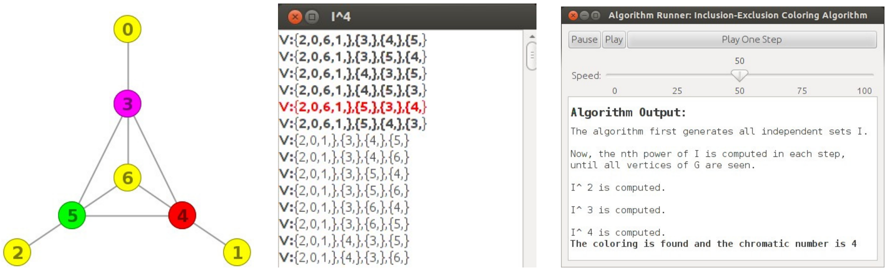



**Fig. 2.** (Left) A minimal coloring of $G$. (Middle) A list of $I^4$ in which one of the elements of $I^4$ is selected. The minimal coloring depicted left corresponds to this row selection. (Right) The final view of the algorithm runner window reporting the chromatic number.

## 4 Concluding Remarks

The focus of this paper is on the educational software GraphTea, which is capable of illustrating algorithms arising from graph theory. We added to GraphTea the new functionality to illustrate the recent graph coloring algorithm that is based on the principle of inclusion and exclusion. The main idea of this algorithm applied to some graph is to repeatedly compute powers of the set of all independent sets of that graph. GraphTea allows the student to interactively explore this algorithm. In particular, it is extremely simple to construct a rich set of different problem instances. This way, the student can easily experiment with different situations and improve his or her understanding of the algorithm. Furthermore, the student can access and visualize the intermediate results that are difficult to inspect without a system like GraphTea. More precisely, the underlying sets of sets of vertices that need to be colored differently are selectable by the student. The corresponding coloring of the vertex sets are automatically available making it easier to understand the intermediate steps of the algorithm.

## References

1. Björklund, A., Husfeldt, T., Koivisto, M.: Set partitioning via inclusion-exclusion. SIAM Journal on Computing 39(2), 546–563 (2009)
2. Bücker, H.M., Rostami, M.A.: Interactively exploring the connection between nested dissection orderings for parallel Cholesky factorization and vertex separators. In: IEEE 28th Int. Par. & Dist. Process. Symp., IPDPS 2014 Workshops, Phoenix, USA, May 19–23, 2014, Los Alamitos, USA, pp. 1122–1129. IEEE Computer Society Press, Los Alamitos (2014)
3. Carbonneaux, Y., Laborde, J.-M., Madani, R.M.: CABRI-Graph: A tool for research and teaching in graph theory. In: Brandenburg, F.J. (ed.) GD 1995. LNCS, vol. 1027, pp. 123–126. Springer, Heidelberg (1996)
4. Lambert, A., Auber, D.: Graph analysis and visualization with Tulip-Python. In: 5th Europ. meeting Python in science, EuroSciPy 2012, Bruxelles, Belgium (2012)
5. Rostami, M.A., Azadi, A., Seydi, M.: GraphTea: Interactive Graph Self-Teaching Tool. In: Proc. 2014 Int. Conf. Edu. & Educat. Technol. II, pp. 48–52 (2014)

# A Demonstration of ALAS-KA: A Learning Analytics Tool for the Khan Academy Platform

José A. Ruipérez-Valiente[1,2], Pedro J. Muñoz-Merino[1], and Carlos Delgado Kloos[1]

[1] Universidad Carlos III de Madrid, Avenida Universidad 30, 28911 Leganés, Madrid, Spain
[2] IMDEA Networks Institute, Av. del Mar Mediterráneo 22, 28918 Leganés, Madrid, Spain
{jruipere,pedmume,cdk}@it.uc3m.es

**Abstract.** Instructors and students have problems monitoring the learning process from low level interactions in on-line courses because it is hard to make sense of raw data. In this paper we present a demonstration of the Add-on of the Learning Analytics Support in the Khan Academy platform (ALAS-KA). Our tool processes the raw data in order to transform it into useful information that can be used by the students and instructors through visualizations. ALAS-KA is an interactive tool that allows teachers and students to select the provided information divided by courses and type of information. The demonstration is illustrated with different examples based on real experiments data.

**Keywords:** learning analytics, Khan Academy, technology-enhanced learning, visualizations.

## 1 Introduction

It is common the use of Virtual Learning Environments (VLEs) to support the learning process. However, it is very hard to interpret the raw data that these platforms generate, thus instructors and students need help to understand the learning process in online courses [1]. For this reason, the transformation of raw data into useful higher level information is required so that low level interactions are not only presented.

Khan Academy is an educational web-site which provides a free education for everyone. The Khan Academy platform is able to provide different types of exercises and videos but also implements some innovative features like gamification techniques and learning analytics visualizations. These default visualizations include low level interactions and also some higher level information. In this work we present a demonstration of ALAS-KA, which is a learning analytics tool that provides new visualizations with new information for Khan Academy which are not available by default. In addition we illustrate this demonstration with real examples of courses where ALAS-KA has been applied, showing visualizations that stakeholders should interpret to understand the learning process. There are other works [2-3] that have developed other similar learning analytics tools for different platforms such as Moodle and WebCT, although the provided information is not the same as the one ALAS-KA provides.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 518–521, 2014.
© Springer International Publishing Switzerland 2014

# 2 Overview of ALAS-KA

## 2.1 Technology

ALAS-KA is designed as a plug-in for the Khan Academy platform. The Khan Academy system as well as ALAS-KA run over the Google App Engine (GAE) architecture and use the GAE Datastore for data persistence. In addition, the underlying programming language is Python. ALAS-KA needs the data generated by students while interacting with Khan Academy to process it to obtain higher level information. Furthermore, we use the Google Charts API for visualizations. An in-depth description of technology aspects have been addressed in previous work [4].

## 2.2 Metrics

It is hard to make sense of raw learning data without prior processing. We have developed a set of metrics that transforms raw data into meaningful information that can be used by the actors who intervene in the learning process. A total set of 20 different parameters have been introduced in ALAS-KA. Specific formulas for some of these metrics have been presented in previous work [5].

We have divided the different metrics into five functional modules [5] which are explained next. *Total Use of the Platform* gives insight about the use students has done in the platform. *Correct Progress on the Platform* contains parameters which try to assess how good the interactions have been. *Time Distribution of the Use of the Platform* focuses on analyzing the distribution of the time in which users have interacted with the platform. *Gamification Habits* offers a couple of metrics to see if students are motivated by gamification elements. Finally, *Exercise Solving Habits* analyzes users' behaviors when solving exercises such as hint avoidance, try abuse or unreflective.

## 2.3 Visualizations

We use visualizations with the aim of presenting the information obtained by the processing of the metrics previously explained. Some of the design criteria followed to implement the interface are the following: keep the interface as simple as possible, make a meaningful use of the colors, divide the user interface in parts and use the same standards in the entire application. The objective is that teachers and students can have a good interaction with ALAS-KA. The implemented visualizations for the explained metrics can be divided in two groups:

- Class visualizations: These visualizations present an overview of the status of the entire class or a set of students. The main type of graphic used for class visualizations are pie charts because we want to give an overview of how the class is distributed for each kind of metric. An example is shown in figure 1.

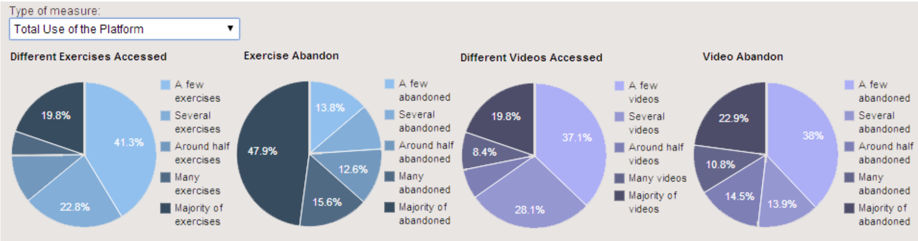



**Fig. 1.** Example of class visualizations retrieved from the *Total Use of the Platform* module

- Individual visualizations: In-depth visualizations enable teachers to analyze each student separately or self-reflection for students. We have decided to use bar charts for individual visualizations. Each type of metric has its individual visualization with two bars: the first bar represents the result for that student and the second the mean value of the class. Two different examples are shown in figures 2-3.

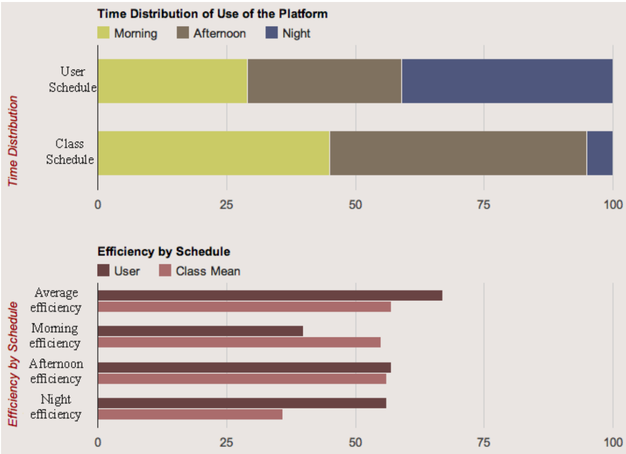



**Fig. 2.** Individual visualizations about the *Time Distribution of the Use of the Platform*

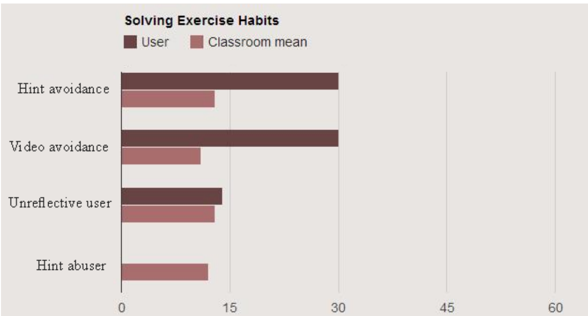



**Fig. 3.** Example of individual visualizations about *Exercise Solving Habits* metrics

## 3 Demonstration Outline

The demonstration will show the different main functionalities of ALAS-KA and how students and teachers can interact with the tool. In addition, the demonstration will show different visualizations in real courses to enable instructors improve their learning decisions and interventions. The live demonstration outline will be as follows:

- The demonstration will start first with a very brief explanation of the objectives of ALAS-KA, as well as its scientific contribution.
- Secondly, the general menus of the tool will be explained (selection of a course, selection of individual or class visualizations, selection of the type of metrics, etc.)
- Next, some class visualizations will be provided using the tool, analyzing the learning process based on these visualizations and promoting the discussion for other possible interpretations depending e.g. on the context.
- Finally, ALAS-KA will be used to present some visualization of individual students in order to make a deep analysis of some selected students. We will give an analysis of these specific visualizations, and the possible interpretations and ways of teachers to act.

**Acknowledgements.** Work partially funded by the EEE project, "Plan Nacional de I+D+I TIN2011-28308-C03-01" and the "Emadrid: Investigación y desarrollo de tecnologías para el e-learning en la Comunidad de Madrid" project (S2009/TIC-1650)".

## References

1. Siemens, G., Long, P.: Penetrating the fog: Analytics in learning and education. In: Educause Review (2011)
2. Mazza, R., Dimitrova, V.: Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education. In: 13th International World Wide Web conference on Alternate Track Papers & Posters, pp. 154–161. EEUU, New York (2004)
3. Zhang, H., Almeroth, K.: Moodog: Tracking Students' Online Learning Activities. Journal of Interactive Learning Research 21(3), 407–429 (2010)
4. Ruipérez-Valiente, J.A., Muñoz-Merino, P.J., Delgado, C.: An architecture for extending the learning analytics support in the Khan Academy framework. In: 1st International Conference on Technological Ecosystem for Enhancing Multiculturality, pp. 277–284. ACM, Salamanca (2013)
5. Muñoz-Merino, P.J., Ruipérez Valiente, J.A., Delgado Kloos, C.: Inferring higher level learning information from low level data for the Khan Academy platform. In: 3rd International Conference on Learning Analytics and Knowledge, pp. 112–116. ACM, Leuven (2013)

# Presentation Trainer: A Toolkit for Learning Non-verbal Public Speaking Skills

Jan Schneider, Dirk Börner, Peter van Rosmalen, and Marcus Specht

Welten Institute, Open University of the Netherlands, Heerlen, The Netherlands
{jan.schneider,dirk.boerner,
peter.vanrosmalen,marcus.specht}@ou.nl

**Abstract.** The paper presents and outlines the demonstration of *Presentation Trainer*, a prototype that works as a public speaking instructor. It tracks and analyses the body posture, movements and voice of the user in order to give instructional feedback on non-verbal communication skills. Besides exploring the background of the feedback theory used by the prototype, this paper describes its conceptual implementation and proposes its use and deployment for the conference demonstration.

**Keywords:** Sensor-based Learning, Immediate Feedback, Open Source, Demonstration.

## 1 Introduction

Interaction modalities in human computer interaction are rapidly growing and getting more intuitive and realistic. Elements of speech and gesture are integrated in commonly available applications, which handle realistic input with devices such as the Microsoft Kinect[1] or various kinds of sensors. The increasing availability of these devices together with their software development kits (SDK), has allowed for the development of new types of learning based platforms with the capacity to track and give feedback about the performance of different activities. The platform presented in this article, called *Presentation Trainer* has the purpose to work as a public speaker instructor. It tracks the user's body posture, movements and voice in order to give feedback of its performance. The software architecture of the platform has been design to allow different implementations of feedback, giving the opportunity to test for the type of feedback most suitable for the user.

## 2 Background

Presentation skills are important in education. However, to acquire presentation skills students need sufficient practice and feedback [1]. *Presentation Trainer* aims to

[1] Microsoft Kinect: `http://www.xbox.com/en-US/kinect`

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 522–525, 2014.
© Springer International Publishing Switzerland 2014

address this. The feedback given by *Presentation Trainer* gives an answer to the question of: "How am I going?" which is one of the three questions for effective feedback presented in [2] by providing information relative to the performance goal of the user. In order to implement an answer to this question we have followed two approaches. The first approach is based on the feedback dimension described as complexity of feedback [3]. The levels of complexity that have been used are the try-again feedback and the correct-response feedback. In case of the try again feedback, the system keeps pointing out that something is being done inappropriately according the predefined rules, until the user corrects it.

In case of the correct response the system tells the user how to adjust. The second approach used for the implementation of the answer to "how am I going?" is based on the elements of feedback loops [4], which are evidence, relevance, consequences and actions. For this first implementation the evidence and relevance are the only elements presented by the platform.

Making the software architecture modular allows the addition and subtraction of different feedback modules giving the opportunity to experiment with different implementations of the answer, finding the one that is suitable for the user allowing it to learn while avoiding some cognitive overload [5].

### 2.1 Prototype Description

The purpose of *Presentation Trainer* is to give immediate feedback of public speaking skills while preparing for a presentation. The core of the platform allows the inclusion of different types of sensors and feedback mechanisms in order to train different skills. At the moment *Presentation Trainer* provides support in developing non-verbal public speaking skills such as body language and voice.

Two different aspects are analysed in the body language: the posture and the movements of the user. Postures such as arms crossed, legs crossed, hands below the hips, hands behind the body and a hunch posture have been defined and are identified by the platform as incorrect postures during a presentation. Once the system detects one of these postures, it sends feedback about it to the user.

Staying still or moving too much during a presentation increases the probability of losing the attention of the audience. Therefore *Presentation Trainer* tracks the user's movements and if they are outside of the predefined threshold feedback is sent.

The current state of *Presentation Trainer* analyses two different elements of the voice. The first element analysed is the speaking cadence, while pauses are important in order to let the audience assimilate what has been spoken, create, curiosity, anticipation and some tension. Long pauses lead to a decrease of attention from the audience; hence the prototype gives feedback when pauses are shorter or longer than the predefined threshold. The second element is the volume of voice if the volume is too high, too low the platform notifies the presenter about it.

*Presentation Trainer* was developed in Processing 2.1[2], which is an open source JAVA-based programing language and development environment. Its OpenGL

[2] Processing: `http://processing.org`

integration allows for fast graphic manipulation and makes it suit for interactive 2D and 3D programs. In order to track the posture and movements from the user, *Presentation Trainer* makes use of the Microsoft Kinect sensor and the OpenNI[3] framework that is an open source SDK used for the development of 3D sensing middleware libraries and applications.

The voice is captured by the computer's microphone and it is analyzed using the Minim[4] audio library that uses the JavaSound API.

The prototype *Presentation Trainer* has 3 different layers: the sensor layer, integration layer and presentation layer. The sensor layer consists in different sensor modules that can be plugged in to the integration layer. Each module is responsible to analyse its data according to presentation skill rules. The integration layer retrieves the data analyses from the different modules in the sensor layer and sends these results to the different modules in the presentation layer, which are the ones responsible to present the feedback to the user.

## 3 Demonstration

In the context of the conference demonstration interested users will be able to use *Presentation Trainer* and test their presentation skills on body posture and movements, voice cadence and volume, and presentation continuity.

The prototype starts tracking the user's body posture, whenever it stands within 1.2 to 3.5m in front of the Microsoft Kinect. Once the tracking starts, the prototype highlights in the screen the limbs of the user that are in a wrong position letting the user know whether its posture is appropriate for doing a presentation. While the user is being tracked the prototype analyses its movements, in the case the amount of movement is outside of the predefined range of movements suitable for a presentation, the prototype suggests the user to move more or to slow down its body movements.

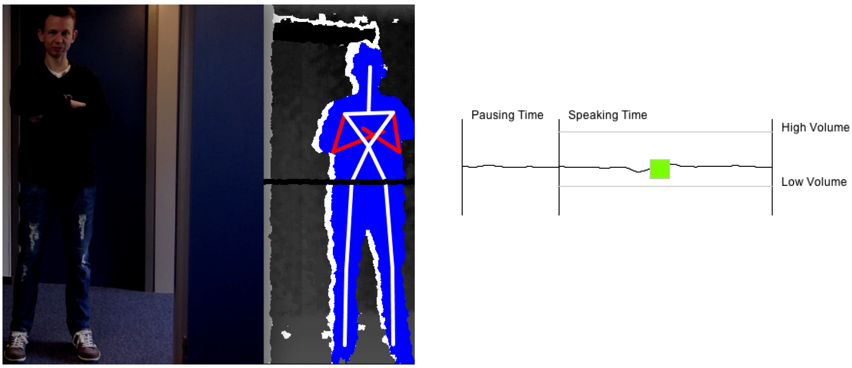



**Fig. 1.** Left: Body Posture Feedback. Right: Voice Feedback

[3] OpenNI: http://www.openni.org

[4] Minim: http://code.compartmental.net/tools/minim/

During the demonstration the user's voice will also be analysed. Suggestions about lowering or increasing the volume of the user's voice are displayed whenever the user's voice volume is outside of the predefined ranges of voice volume suited for giving a presentation. The prototype also reminds the user to start speaking in case it has been quiet for too long time, and suggest the user to take a deep breath and pause whenever it has been talking without pauses for a long period of time.

## 4 Future Work

The planning for experiments testing the feasibility and usability of the presentation trainer is in current development. After the first experiments there are plans to upgrade the presentation trainer into a system able to adapt its feedback on running time according to the performance of the user. For this a set of presentation tasks will be designed with varying difficulty and different levels of support and feedback.

The knowledge gained from *Presentation Trainer* and its immediate feedback will be used for the instructional design in the Metalogue project (www.metalogue.eu), which purpose is to train people for negotiation dialogs, by giving them metacognitive feedback.

**Acknowledgement.** The underlying research project is partly funded by the METALOGUE project. METALOGUE is a Seventh Framework Programme collaborative project funded by the European Commission, grant agreement number: 611073 (http://www.metalogue.eu).

## References

1. Kerby, D., Romine, J.: Develop Oral Presentation Skills Through Accounting Curriculum Design and Course-Embedded Assessment. Journal of Education for Business 85(3) (2009)
2. Hattie, J., Timperley, H.: The power of feedback. In: Review of Educational Research, pp. 81–112 (2007)
3. Mory, E.H.: Feedback research revisited. In: Handbook of research on educational communications, pp. 745–784 (2004)
4. Goetz, T.: Harnessing the power of feedback loops (2011), `http://www.wired.com/2011/06/ff_feedbackloop/all/` (accessed on april 2014)
5. Jong, T.: Cognitive load theory, educational research, and instructional design: some food for thought. Instructional Science 38(2), 105–134 (2009)

# Designing Visualisation and Interaction for Social E-Learning: A Case Study in Topolor 2

Lei Shi and Alexandra I. Cristea

Department of Computer Science, University of Warwick, CV4 7AL, Coventry, UK
{lei.shi,acristea}@dcs.warwick.ac.uk

**Abstract.** This paper presents the design of the visualisation and interaction in Topolor 2, an online learning system that supports social e-learning using adaptation and personalisation techniques. The main goal of this study is to investigate if and how the visualisation and interaction design based on self-determination theory (SDT) can improve the learning experience and learning efficiency. The goal of this paper and demonstration is to showcase the new design of the visualisation and interaction features hands-on at the conference.

**Keywords:** self-determined theory, social e-learning, system design, evaluation.

## 1 Introduction

Social e-learning is a process in which learners achieve learning goals by sharing knowledge, skills and abilities in a community where they have common interests [5]. Sustainable social e-learning occurs along with learners' interaction with one another and their motivation in participation. The abundant resources shared within a learning community offer opportunities for learners to connect to various learning content and peers, whereas one of the biggest challenges is how to direct and motivate learners to access the most appropriate resources and benefit from them. Adaptation and personalisation have been suggested as potential solutions to tackle this challenge [3].

Topolor [4] is an online learning system that supports social e-learning using adaptation and personalisation techniques, based on the hypothesis that *extensive social features, personalised recommendations and Facebook-like appearance would make a system more familiar to the learners, and subsequently increase learning experience*. This paper demonstrates the design of the novel visualisation and interaction features in Topolor 2, which have been newly introduced, based on a solid theoretical foundation, the self-determination theory (SDT) [2]. These new features aim at fulfilling the learners' perceived psychological needs, so as to direct and motivate them to access learning resources in a social e-learning context. The ultimate goal is to investigate if and how the visualisation and interaction features designed based on self-determination theory (SDT) can improve the learning experience and learning efficiency. The goal of this paper and demonstration is to briefly introduce these new features, as well as give researchers a hands-on experience of them at the conference, and gather feedback from peers and e-learning specialists.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 526–529, 2014.
© Springer International Publishing Switzerland 2014

## 2 Self-Determination Theory

Self-determination theory (SDT) is an empirical motivation theory, which proffers that individuals become increasingly self-determined and self-motivated when they feel satisfaction of their psychological need of autonomy, competence and relatedness [2]. According to SDT, people strive for *autonomy* over their actions and decisions; they tend to obtain *competence* in their actions and surroundings; and since activities such as learning often occur in a social context, *relatedness* is proposed as the third essential psychological need. In all, SDT defines the following needs [2] to fulfil:

- *Autonomy*: a sense of internal assent of one's own behaviours;
- *Competence*: a sense of controlling the outcome and experience mastery;
- *Relatedness*: a sense of connection and interaction with others within a community.

It is expected that an e-learning system that fulfils all the above three psychological needs can sustainably increase the learners' intrinsic motivation, and thus improve the learning experience and learning efficiency [1]. Therefore, in this study, SDT guides the design of the visualisation and interaction in the second version of Topolor.

## 3 Visualisation and Interaction Design Based on SDT

**To Fulfil the *Autonomy* Need.** Topolor lets learners feel in charge of their own behaviours and feel in control of their learning progress. A learner can clearly see the 'layered' learning goal of the course, i.e., the long-term goal to complete the course, the medium-term goal to finish a topic and the short-term goal to read a resource, as well as the reason for recommendations. A learner knows what is happening in the community on the home page (Fig. 1 (a)), a course page (Fig. 1 (c)) and a topic page (Fig. 1 (f)) that show various lists of, e.g., resources and peers' activities, and she can adjust the lists using various filter-sorter widgets (Fig. 1 (a. 2) and (c. 2))).

**To Fulfil the *Competence* Need.** Topolor lets learners feel mastery of their skills and confidence in current context, where cognition and expectation are consistent with system responses, so as to be able to obtain further skills and confidence with relative ease. A learner can see her learning progress in various ways. For example, in the learning path pop-up view (pops up when clicking on the button 'Learning Path' on a course page, as shown in Fig. 1 (c), or on a topic page, as shown in Fig. 1 (f)), a learner can see which topics she hasn't learnt, which topics she is eligible to learn, and which topic is recommended for her to learn, as shown in Fig. 1 (e). A learner can also see her performance on a course (by clicking on the button 'My Performance' on a course page, as shown in Fig. 1 (c)), and on a topic (by clicking on the same button on a topic page, as shown in Fig. 1 (f)), in a pop-up view that shows her performance, such as the number of courses/topics she has learnt via a line-chart (shows when clicking on the tab-button 'course/text', as shown in Fig. 2 (h)). To strengthen the feeling of competence need satisfaction, Topolor provides some enjoyable and fun features, such as the one shown in Fig. 1 (g) that lets a learner compare ones own contributions to others ('PK.': 'Player Killer', naming convention taken from games).

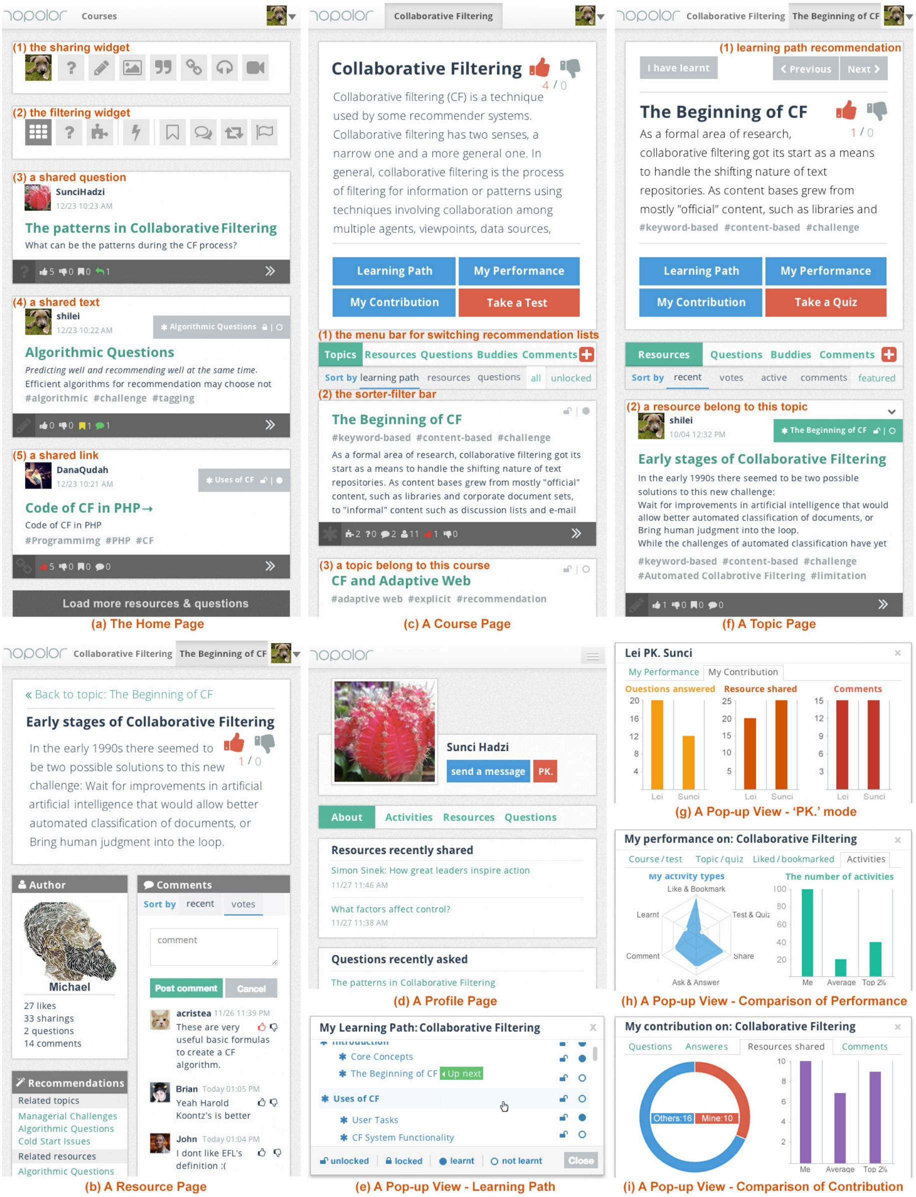



**Fig. 1.** The user interface of Topolor 2

**To Fulfil the *Relatedness* Need.** Topolor lets learners feel connected to peers in the social learning community by supporting social status visualisation and comparative social visualisation. A learner can view her or other 'social statuses' on a question or a resource page's author information panel (Fig. 1 (b), showing statistic numbers), and a peer's profile page (Fig. 1 (d), showing a learner's learning activities). A learner can check her contribution to the learning community and compare her performance and contribution with others. The data in the user model are visualised and opened to the learners (Fig. 2 (g), (h), (i)), which potentially encourages them to contribute more to the community, as seeing each other's status may stimulate imitation and competition.

## 4 Conclusion and Future Work

This paper has presented the new visualisation and interaction features in Topolor 2, an online social e-learning system using adaptation and personalisation techniques. These new features were designed rooted in self-determination theory (SDT), aiming at improving the learning experience and learning efficiency in social e-learning. We intend to demonstrate Topolor 2 to researchers and practitioners in the conference and both showcase these new features and gather feedback on further improvements.

The preliminary evaluations were conducted using questionnaires that consisted of seven 5-point Likert scale (from 1: *strongly disagree* to 5: *strongly agree*) questions asking the comparison between Topolor 2 and other regular e-learning systems, including *I believe Topolor* ***a)*** *helped me learn more topics;* ***b)*** *helped me learn more profoundly (deeply);* ***c)*** *increased my learning outcomes more;* ***d)*** *was easier to use;* ***e)*** *was more useful;* ***f)*** *was easier to learn how to use; and* ***g)*** *was easier to remember how to use*. 25 completed questionnaires were collected. The mean values for each of the answers were ***a****: 3.6*, ***b****: 4.0*, ***c****: 4.0*, ***d****: 4.0*, ***e****: 4.0*, ***f****: 4.1* and ***g****: 4.3* (all>3.0, the neutral value). The standard deviation values for each of the answers were ***a****: 0.71*, ***b****: 0.76*, ***c****: 0.68*, ***d****: 0.57*, ***e****: 0.64*, ***f****: 0.67* and ***g****: 0.74*, and the results' *Cronbach's alpha* value was 0.82 (≥0.8, indicating the results were reliable). The results indicated that the visualisation and interaction designed based on SDT had positive effect on the perceived learning experience and learning efficiency. The future work will start with investigating how each feature can improve the learning experience and learning efficiency, analysing the usage data tracked by Topolor's logging mechanism.

## References

1. Muntean, C.I.: Raising engagement in e-learning through gamification. In: Proc. of the 6th Intl. Conf. on Virtual Learning, pp. 323–329 (2011)
2. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55(1), 68 (2000)
3. Shi, L., Cristea, A.I., Foss, J.G.K., Al-Qudah, D.A., Qaffas, A.: A social personalized adaptive e-learning environment: a case study in Topolor. IADIS Int. J. WWWInternet. 11(2) (2013)
4. Shi, L., Al-Qudah, D.A., Cristea, A.I.: Designing Social Personalized Adaptive e-Learning. In: Proc. of the 18th ACM Conf. on Innovation and Technology in Computer Science Education, p. 341 (2013)
5. Shi, L., Al-Qudah, D.A., Qaffas, A., Cristea, A.I.: Social e-learning in Topolor: a case study. In: Proc. of the 7th IADIS Conf. E-Learning, pp. 57–64 (2013)

# GPIM: Google Glassware for Inquiry-Based Learning

Angel Suarez, Stefaan Ternier, Marco Kalz, and Marcus Specht

Welten Institute, Research Centre for Learning, Teaching and Technology
Open University of the Netherlands; Heerlen, The Netherlands
{angel.suarez,stefaan.ternier,marco.kalz,marcus.specht}ou.nl

**Abstract.** Over the past few years, the use of mobile personal devices has witnessed a widespread take-up. With wearable technology like head-up-displays a new genre of educational technology is appearing to enhance contextualized learning. This paper reports about a Google Glass prototype for Inquiry-Based Learning (IBL). With the use of Google Glass we aim to bring learning to the next level providing a more seamless experience where Glass supports the learning- and inquiry-process just in time and in an unobtrusive way. This demo paper introduces the design and functionalities of Google Glass for the Personal Inquiry Manager (GPIM). The paper concludes with open issues for future research, especially focused on evaluation and future development.

## 1 GPIM Implementation

The weSPOT project [1] [2] aims to foster scientific inquiry as the approach for science learning and teaching, linking everyday life with science teaching in schools using technology. weSPOT supports scientific concepts by relating them with personal curiosity, experiences and reasoning. The project has three main goals: implement a working environment that allows the easy linking of inquiry activities with school curricula and legacy systems, create a diagnostic instrument for measuring inquiry skills and work out a reference model to foster Inquiry-Based Learning (IBL) skills. This model consist of 6 phases:

Students or learners formulate *questions and hypotheses* about issues that they wonder about. In the *operationalization* phase learners define the concepts within their topics that are unclear and make them distinguishable or measurable, in order to allow the empirical observations that will take place in the data collection process. In the *data collection* phase, learners collect evidences for their inquiries by taking pictures, recording videos or audio recordings. Once the data has been collected, learners are able to analyze these data in the *data analysis* phase. Data is treated and exploited to derive conclusions, extracting useful information or supporting decision-making. In cases where an experiment has been performed, a statistical analysis can be required. *Interpretation* or *discussion* describes the relevance of the results confirming or rejecting the questions that have been posed before. This discussion should relate the obtained conclusions to the existing body of research, suggest where current assumptions may be modified because of new evidence, and possibly identify

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 530–533, 2014.
© Springer International Publishing Switzerland 2014

unanswered questions for further research. The last phase of the process is *communication*, where results, findings and conclusions are shared or published with colleges or stakeholders who may use these results afterwards.

One aspect of this project is the development of mobile applications to support and structure the inquiry-process. To offer new interaction possibilities with the weSPOT inquiry environment we have developed the Google Glass Personal Inquiry Manager (GPIM). The GPIM (Glass PIM) is Glassware (a native application for Google Glass) to support the inquiry based learning process. The application supports currently three out of the six phases of the IBL model: the moment of curiosity that students can have or *wonder moment*, the question/hypothesis phase and the data collection phase. Figure 1 shows the screens that guide the user through these phases.

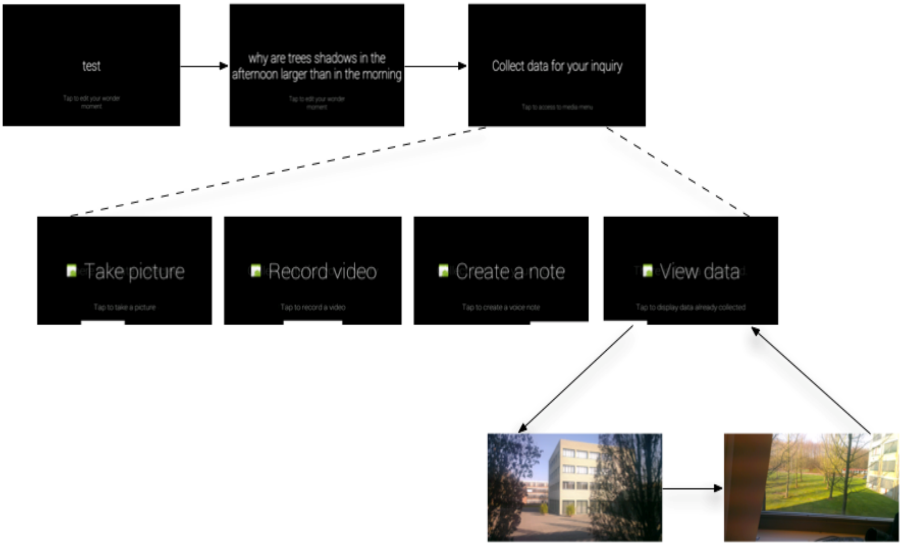



**Fig. 1.** GPIM workflow

The user starts the GPIM to support his or her inquiry process. To activate the GPIM the student will not need to interact physically with the device. Just in time and in an unobtrusive way students are able to use their voice to launch the first screen. The voice command “Ok Glass, new inquiry” is a trigger that starts a new inquiry.

During this initial phase, the user makes a statement about what he or she wonders about. For instance, a student notices that the shadows of the trees are larger in the afternoon than in the morning. This simple idea can trigger an entire spontaneous inquiry about the rotation and translation of the sun.

From this initial screen, the users can formulate their *wonder moment*. The GPIM supports capturing the wonder moment via voice interaction. As the user taps the GPIM in this context, glass will start recording the student’s voice and via voice recognition capture this moment of curiosity.

Now that wonder moment has been set, the inquiry process can begin. Learners need to come up with a hypothesis that they would like to confirm or reject running the inquiry. Therefore the next screen will allow them to set up one or more hypotheses. The inquiry process is not tied to a specific question, it should enable open research about multiple questions that can be related but they are not necessary the same. For instance, following the previous example, a student can set up these two hypotheses: (1) there is a relation between the shadow's length and the position of the sun, and (2) therefore it is possible to calculate the height of big objects, such as buildings, that can not be measured easily because their dimensions.

The next step of the inquiry workflow aimed to collect evidences in order to confirm or reject the assumptions and hypotheses made. Current step enables take pictures, record videos or create a voice notes (Figure 1). After data has been collected, the number of items collected is shown and a new option is available in the menu to display the evidences collected.

Although the main goal of GPIM is not displaying the data collected, the student can visualize his or her data using *View Data* option. This option presents the data collected in a timeline (Figure 1), where old data is placed at the end of the timeline and recent data will appear at the beginning. In addition, the student is able to scroll between data collected using the touchpad. Furthermore, every single item has its own menu in order to provide specific functionality such as play or resume for audio and video media.

The GPIM synchronizes data with the server when the user is online. However, the nature of an inquiry will lead to users stepping in and out of zones of wireless internet. When no internet is available, the GPIM will continue to work in an offline mode, enabling the users to continue capturing and viewing data. This makes the GPIM suitable for every context, including areas without internet connection. In addition, an inquiry can be an ongoing task. With the GPIM, inquiries spanning several days can be paused and resumed whenever the student wants.

## 2 Discussion and Future Work

This article presents Google Glass for the PIM (GPIM), which aims to bring learning to the next level providing a more seamless experience using an innovative and wearable head-up display technology. Particularly, what GPIM offers is innovation where technology adapts to the learner's habits of learning offering a first view perspective and enabling fully interaction with the environment. Using head-up displays for inquiry-based learning can be especially useful in situations in which learners need to document something while using their hands. Experimentation phases in a chemistry lab are a good example for inquiries that are hard to capture with traditional approaches or they need a high amount of preparation.

The authors plan further experimentation and optimization of the Glassware PIM with experiments in the following topics:

1. One line of further development will focus on the specific added value of the GPIM. Therefore the authors will create different predefined inquiries and

learners will use the GPIM to work on different activities in the inquiry phases, which might be appropriate for Glassware. The goal will be to find the optimal fit for using Google Glass based applications in a whole set of devices and inquiry activities.
2. A second line of research will especially look at differences of the results of data collected by Google Glass GPIM compared to data collected with smartphones and tablets as we expect quality differences in these.

These experiments are intended to provide outcomes for the next research questions:

1. What is the impact on student's motivation using GPIM instead of PIM?
2. What is the impact on perceived flexibility and usability using GPIM?
3. Can mobile technology and data collection enable the integration of "moments of curiosity" (wonder moments) in the IBL process?

Besides the main goal of this manuscript, which has been introduce the GPIM Glassware as a way of supporting the IBL process, it also has been intended to contribute with a new branch for research to discuss the impact of wearable HUD technologies like, Google Glass in learning environments.

**Acknowledgments.** The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement N° 318499 - weSPOT project.

## References

1. Specht, M., Bedek, M., Duval, E., Held, P., Okada, A., Stefanov, K., Parodi, E., Kikis-Papadakis, K.: Strahovnik: weSPOT: Inquiry Based Learning Meets Learning Analytics. In: The Third International Conference on e-Learning, Belgrado, Serbia (2013)
2. Mikroyannidis, A., Okada, A., Scott, P., Rusman, E., Specht, M., Stefanov, K., Boytchev, P., Protopsaltis, A., Held, P., Hetzner, S., Kikis-papadakis, K., Chaimala, F.: Wespot: A Personal and Social Approach to Inquiry-Based Learning. J. Univers. Comput. Sci. 19, 2093–2111 (2013)

# Lifelong Learning Hub: A Seamless Tracking Tool for Mobile Learning

Bernardo Tabuenca, Marco Kalz, and Marcus Specht

Welten Institute, Research Centre for Learning, Teaching and Technology
Open University of the Netherlands; Heerlen, The Netherlands
{bernardo.tabuenca,marco.kalz,marcus.specht}ou.nl

**Abstract.** Lifelong learners' learning activities are scattered along the day in different locations and they make use of multiple devices. Most of the times adults have to merge learning, work and everyday life making it difficult to have an account on how much time is devoted to learning activities and learning goals. Learning experiences are disrupted and mobile seamless learning technology has to find new solutions to integrate daily life activities and learning in the same process. Hence, there is a need to provide tools that are smoothly integrated into adults' daily life. The contribution of this demonstration is presenting a mobile tool that leads the lifelong learning towards a self-regulated process: foster awareness on learning goals and learning moments; facilitates the user to keep track of learning time with frictionless interface; fosters engagement and motivation on the task providing useful statistics. The 3LHub project has been released under open access with the aim to foster adaptation to further communities as well as to facilitate the extension to the increasing number of NFC tags existent in the market.

**Keywords:** seamless, lifelong learning, feedback, seamless learning, self-regulation, natural interaction, NFC, sensors.

## 1 Introduction

Nowadays, lifelong learners are confronted with a broad range of activities they have to manage everyday. In most cases they have to combine learning activities and their professional and private life linking formal and non-formal learning activities. In the setting of an adult lifelong learner this is especially difficult as in most cases interests might be highly distributed over different domains and keeping up learning needs an extra effort. One of the main challenges here is the bridging of learning activities between different contexts.

Mobile seamless learning technology can offer solutions to the address this problem [1] [2] The learner-centric view of mobile seamless learning [3] suggests that *a seamless learner should be able to explore, identify and seize boundless latent opportunities that his daily living spaces may offer to him (mediated by technology), rather than always being inhibited by externally-defined learning goals and resources*.

In summary, there is little support for lifelong learners that typically try to learn in different contexts, are busy with multiple parallel learning tracks, and must align or

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 534–537, 2014.
© Springer International Publishing Switzerland 2014

relate their learning activities to everyday leisure and working activities. Candy [4] has summarized four components of self-directed lifelong learning. These are *self-monitoring*, *self-awareness, self-management* (planning of learning) and last but not least *meta-learning*. To date, there is little technological support to enable learners in conducting these different activities across contexts and locations. Hence, there is a need to provide suitable tools for lifelong learners to facilitate bridging learning experiences in a seamless flow. In this paper Near Field Communication (NFC) is proposed as an instantiation for natural interaction with mobile devices and for seam- less integration of technology in lifelong learning.

## 2 The Lifelong Learning Hub

The *LifeLong Learning* Hub[1] (3LHub) is a standalone application developed for NFC-enabled Android (4.03 or above) devices released in March 2014 in Google Play. 3LHub is a mobile seamless tool for self-regulated learning that aims to cover the following gaps in lifelong learners' learning process:

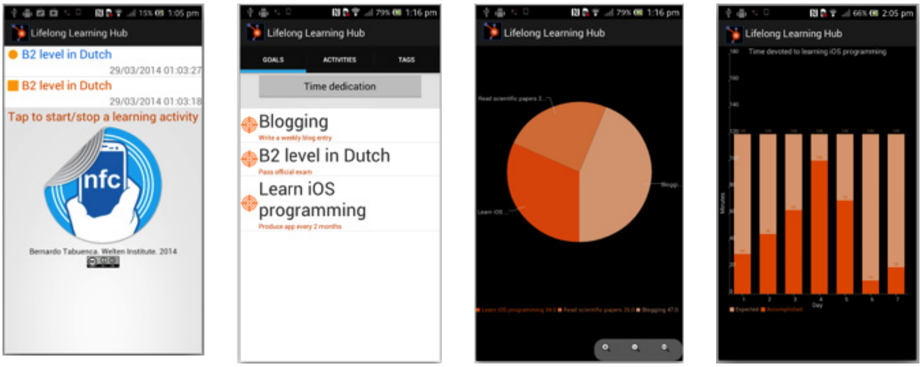
1. No support for learning activities across locations, devices and environments.
2. No linking between learning activities and everyday life.
3. Incompatibility between NFC-tags and NFC-readers is blocking further expansion of the NFC technology for open (educational) communities.
4. No feedback on lifelong learning activities.

The 3LHub has been designed based on the seamless notion that lifelong learners can learn in a variety of scenarios and can switch from one scenario or context to another easily and quickly, using the personal device as a mediator. This tool has been conceptualized on the idea that mobile technology can be smoothly integrated in daily life activities whenever interacting with it requires the least number of clicks (zero) possible and the duration of any action with the tool lasts not longer than 20 seconds. The 3LHub features the following functionalities:

### 2.1 Set Goals

This stage assumes that the user reflects on his autobiography as a learner mapping learning goals to learning environments identified by NFC tags (See use cases in figure 2). Whenever the user sets a new goal in 3LHub, he should get a NFC-tag, tap it with the NFC-enabled mobile device, characterize the goal with a name, specify the expected outcome when the goal is finished, estimate how much time (in minutes) will he devote to this goal on daily basis, and indicate when the goal should be finished. Placing an NFC-tag in a physical learning environment enables the connection of a variety of tracking data with the learning activity. For example the "check-in" at a NFC tag can track the learners use of a specific resource, at a certain time of the day, in a specific location.

[1] The Lifelong Learning Hub project.
`https://sites.google.com/site/lifelonglearninghubproject/`




(a) Tap to start/stop a learning activity

b) Learning goals configured by the user

(c) Overall learning-time by goals

(d) Learning time estimated VS invested for a goal in a week

**Fig. 1.** Binding goals to NFC-tagged learning environments with 3LHub

## 2.2 Perform/Track Learning Activity

Lifelong learners recur to specific locations (e.g. desktop, coach) and moments (e.g. waiting times, transitions) to accomplish their learning activities along the week [5]. Learning activities should be tracked in a way that the transition from/to daily life activity can be done with effortless interaction, otherwise the user will not bother to track such a short learning moments, and as result it will never be accounted as learning time.

a) Write a blog entry at coffee time when I arrive to work

b) Reading a pocket-book in waiting times

c) Listening foreign language podcasts in car transitions

d) Watch MOOC videos in tablet during commercial-breaks

**Fig. 2.** Learning activities (write, read, listen, watch) bound to daily learning environments

This feature considers that the user will tap the associated NFC-tag every time he starts/stops a learning activity (Figure 1a). Hence, 3LHub harvests all learning moments and accounts them as real learning time with frictionless interactions.

### 2.3 Monitor Learning Activities

The 3LHub features the following visualizations with the aim to foster understanding on learning habits, optimise learning, and, bind successful learning environments:

1. *Percentage of time invested on each learning goal*. Figure 1c illustrates how percentage of total time and number of minutes are presented in a pie chart. This visualization can be used by lifelong learners to compare time invested on his learning goals, identify priorities to accomplish goals, and, patterns regarding preferences for specific learning environments, devices or learning activities (read, watch, write, listen) Fig. 2.
2. *Distribution of learning moments along the day*. This feature illustrates the distribution of the learning moments during the day (X axis 0..24) for a whole week (Y axis 1..7).
3. *Monitoring accomplished goals*. Figure 1d illustrates a representation of accomplished learning time versus expected time towards a learning goal 3LHub.

## 3 Conclusions

The contribution of this demonstration is presenting a tool for lifelong learners to bridge scattered personal learning environments in which learners can define their personal ecosystem and experience the interaction with such a system in long term typical lifelong learner settings. This research aims at giving an open, flexible and low-cost prototyping framework for defining and linking everyday learning activities to contexts, physical artefacts, everyday home media solutions, and supporting to link sustainable learner tracks to these components. In future research, this tool will be evaluated in a longitudinal study with the aim to estimate how accurate are the estimations made by the teachers in contrast to real measures made by the students.

## References

1. Chan, T.-W., Roschelle, J., Hsi, S., Sharples, M., Brown, T., Patton, C., Cherniavsky, J., Pea, R., Norris, C., Soloway, E., Balacheff, N., Scardamalia, M., Dillenbourg, P., Looi, C.-K., Milrad, M., Hoppe, U.: One-To-One Technology-Enhanced Learning: an Opportunity for Global Research Collaboration. Res. Pract. Technol. Enhanc. Learn. 01, 3–29 (2006)
2. Wong, L.-H., Looi, C.-K.: What seams do we remove in mobile-assisted seamless learning? A critical review of the literature. Comput. Educ. 57, 2364–2381 (2011)
3. Wong, L.-H.: A learner-centric view of mobile seamless learning. Br. J. Educ. Technol. 43, E19–E23 (2012)
4. Candy, P., Brookfield, S.: Self-direction for lifelong learning: A comprehensive guide to theory and practice, San Francisco, USA (1991)
5. Tabuenca, B., Ternier, S., Specht, M.: Supporting lifelong learners to build personal learning ecologies in daily physical spaces. Int. J. Mob. Learn. Organ. 7, 177–196 (2013)

# Adventure-Style Serious Game for a Science Lab

Vasilis Zafeiropoulos, Dimitris Kalles, Argyro Sgourou, and Achilleas Kameas

Hellenic Open University, Patras, Greece
{vasiliszaf,kalles,sgourou,kameas}@eap.gr

**Abstract.** Universities and medical companies deal constantly with the issue of training their students and staff in making proper use of their laboratories. Laboratory skills and accurate implementation of safety rules cannot be developed through simple observation. Practice in a laboratory requires dedication, time and space flexibility and controlled experimentation with the highly sensitive and expensive lab equipment. Specifically, hands-on laboratory workshops or "wet workshops" combine theoretical and practical training and perform step by step experimental procedures including various lab instruments and potential hazardous substances to educate participants. To overcome such limitations and provide both realism and user-friendliness, Onlabs demo, has been developed to reflect a modern commercial 3D Adventure Game; the user uses the arrow keys to navigate in the lab and the mouse to press buttons, turn knobs, use specific objects alone or in pairs and collect several of them in their inventory. Onlabs simulates the biology laboratory at the Hellenic Open University, Greece and already fulfills a detailed overview of the photonic microscopy and the preparation of buffer solutions.

**Keywords:** Serious Games, Virtual Worlds, Biology Laboratory, Distance Learning.

## 1 Introduction

Open Universities and medical companies face the challenge to establish distance educational laboratory courses and familiarize students and employees respectively to make proper use of their laboratory equipment and to conduct experiments successfully and with safety. As the laboratory equipment is, in most cases, sensitive and expensive, the trainees are not allowed to learn by “trial-and-error” – rather they are strictly instructed what to do; laboratory training should comply with all safety rules. Moreover, within laboratory areas, only a specific, usually low, number of trainees is permitted as to avoid potential accidents or damages to the lab equipment and all participants are obliged to follow instructions from their supervisors. As a consequence, it turns into a stiff learning process, while the learning outcome cannot be expected to be the optimal. As an alternative option, a computer game providing the trainees with a simulation environment of the lab and a user-friendly interface for their interaction with it would be a helpful tool for their laboratory training before they proceed to the hands-on experimental practice. Our demo, Onlabs, is a 3D adventure-like game which aims

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 538–541, 2014.
© Springer International Publishing Switzerland 2014

to overcome the afore-mentioned obstacles; in brief, it consists of a virtual environment simulating the biology laboratory at the Hellenic Open University (HOU) where a human user interacts with the laboratory equipment and acquires familiarity with it and its use, and to some extent, can successfully conduct experiments.

## 2 Aims

Learning through virtual classes is known to be more powerful than through lectures or books; reading books and attending lectures help students learn a subject by 10% or even lower but interacting through a virtual world increases the learning rate to 75% [1].

Speaking specifically of biology laboratory training, we expect that an immersive virtual laboratory class, provided with a user-friendly interface, could be an efficient tool for teaching university students and medical companies' employees how to make proper use of the on-site laboratory.

## 3 Implementation

Onlabs software uses the Hive3D Development Platform, by Eyelead Software, a company based in Athens, Greece.

### 3.1 Control

To provide both realism and user-friendliness, the interface between the human agent (user) and the environment has been developed in a way that it resembles the one of a modern commercial 3D Adventure Game; the user uses the arrow keys to navigate in the lab while they use the mouse to press buttons, turn knobs, use specific objects alone or in pairs and collect several of them in their inventory so as to make a combined use of the objects and follow a sequential series of actions of an experimental protocol.

### 3.2 Scenario of the Game

The main instrument used by the user in the game is the microscope. Specifically, they can connect it to and disconnect it from the socket, turn on or off the light switch and turn left or right the knobs configuring the light intensity, the stage's height, the ocular lenses' position etc.

The game scenario or, in other words, the necessary sequence of steps that the user must follow in order to accomplish their "mission", consists of the basic steps required for the preparation of the microscope as they are described in the manual of the (on-site) biology lab at HOU: connecting the microscope to the socket; turning on the light switch; configuring the light intensity at the ¾ of the maximum value; lifting the condenser knob at the highest position; rotating the rotating knob with the objective lenses of the microscope so that the 4X lens (the one with the lowest magnification), becomes active; testing the coarse focus knob by rotating it right and left; and finally testing the fine focus knob by rotating in right and left.

Apart from those necessary steps, the user is also given the chance to create a testing concoction by successively combining in pairs several objects such as a piece of paper with the letter "E" written on it, a nosepiece, a cover glass, a pipette and a beaker full of water, and finally microscope that concoction with each one of the 4 objective lenses, achieving different focuses through the rotation of the various knobs of the microscope.

**Fig. 1.** Graphical design of the biology laboratory's microscope and equipment at Hellenic Open University

## 4 Evaluation

Onlabs was evaluated by 19 biology students, whose ages ranged from 35 to 50 and all of them were member of the same class.

The evaluation was based on questions chosen from System Usability Questionnaire (SUS) for systems engineering, developed by John Brooke of Digital Equipment Corporation in Great Britain, in 1986 [2].

The evaluation results were encouraging; all students found the application convenient to use and suggested it had been as a helpful learning tool for them.

## 5 Conclusion

Onlabs is a 3D game for training HOU's biology students in appropriately using the biology lab and successfully conducting experiments. A class of 19 biology students at HOU tested it and evaluated it as a user-friendly and helpful learning tool.

A demo version of Onlabs is available for downloading and playing [3].

## 6 Future Work

Onlabs is already being expanded in order to gradually include all of HOU's laboratory instruments; moreover, the game interface is being expanded in order to include more actions and improved in terms of user-friendliness.

Longer-term goals include the investigation of interaction between virtual and human agents, the deployment of machine learning for training virtual tutors to oversee online human students and the exploitation of an institutional incentive on specifying the learning outcomes of study modules [4] as a means for reaching out to vocational training and high school audiences whose courses are closely aligned with our science labs learning content.

**Acknowledgements.** Our work takes place within a university project called Onlabs revolving around Vasilis Zafeiropoulos's doctorate "Computer-Human Learning Interaction in Laboratory Environments". "This project has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) (Funding Program: "HOU"). We would also like to express our gratitude to the graphics designer Socratis Galiatsatos, who supplied us with state-of-the-art 3D graphics for Onlabs.

## References

1. De Freitas, S.: Serious Virtual Worlds – A scoping study. JISC (2008)
2. Brooke, J.: SUS: A quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, A.L. (eds.) Usability Evaluation in Industry, Taylor and Francis (1996)
3. Onlabs @ Hellenic Open University, `https://sites.google.com/site/onlabseap/`
4. Aggelopoulou, N., Pierrakeas, C., Artikis, A., Kalles, D.: Ontological Modelling for Intelligent e-Learning. In: 14th IEEE International Conference on Advanced Learning Technologies, Athens, Greece (2014)

# A Multi-purpose Ecosystem for Systematic Gesture Control in Educational Settings

Raphael Zender, Marlene Karlapp, Fritz Rose, David Silva-Ribeaux, Volker Wölfert, and Ulrike Lucke

Department of Computer Science, University of Potsdam, Potsdam, Germany
forename.surame@uni-potsdam.de

**Abstract.** Today, the market on Natural User Interfaces (NUI) is dominated by numerous products that allow the recognition of a broad spectrum of body gestures. On the one side, those technologies change how we use IT and promise new ways to interact with each other via IT. Innovative communications settings appear, for instance for educational institutions that break new grounds beside the traditional on-site teaching and learning. But on the other side, a field of technologies arises that is fragmented by heterogeneous solutions that make an effective application without technology lock-in difficult.

The work behind this demo paper motivates the use of gesture-related NUI systems for educational settings and addresses the integration of heterogenous NUI systems as one of the biggest current challenges. A systematic integration approach including exemplary implementations for lecturers will be introduced.

## 1 Motivation

In the last few years, gestures at whole became more and more relevant as corresponding hardware becomes affordable and reliable. Today's gestural systems enhance the human control on IT, the feeling of control and empowerment, the convenience, and the delight [1]. Unfortunately, different systems make use of different conventions, and there is not yet a standard set of interfaces or libraries.

Today's innovative NUI technologies are inexpensive to buy, address a large target group and come with open Application Programming Interfaces (APIs) for software developers. A broad spectrum of those technologies are based on prototypes without an experienced software forge behind them that considers interoperability with other technologies. Therefore, many of the resulting applications are isolated and monolithic. To integrate the broad spectrum of standalone NUI tools into a systematic architectural paradigm, generic integration strategies are required.

This work introduces a systematic integration approach including one exemplary implementation for one gesture control application in an educational setting.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 542–545, 2014.
© Springer International Publishing Switzerland 2014

## 2 Integrative NUI Ecosystem

The proposed integrative NUI ecosystems strictly separates gesture recognition tools and technologies like the Kinect from applications like visualizations or even controllable physical systems. These two ends of a gesture-controlled IT system are brought together by a service layer that abstracts from technological details in terms of a Service-oriented Architecture (SOA). Figure 1 sketches the general 3-layer architecture.

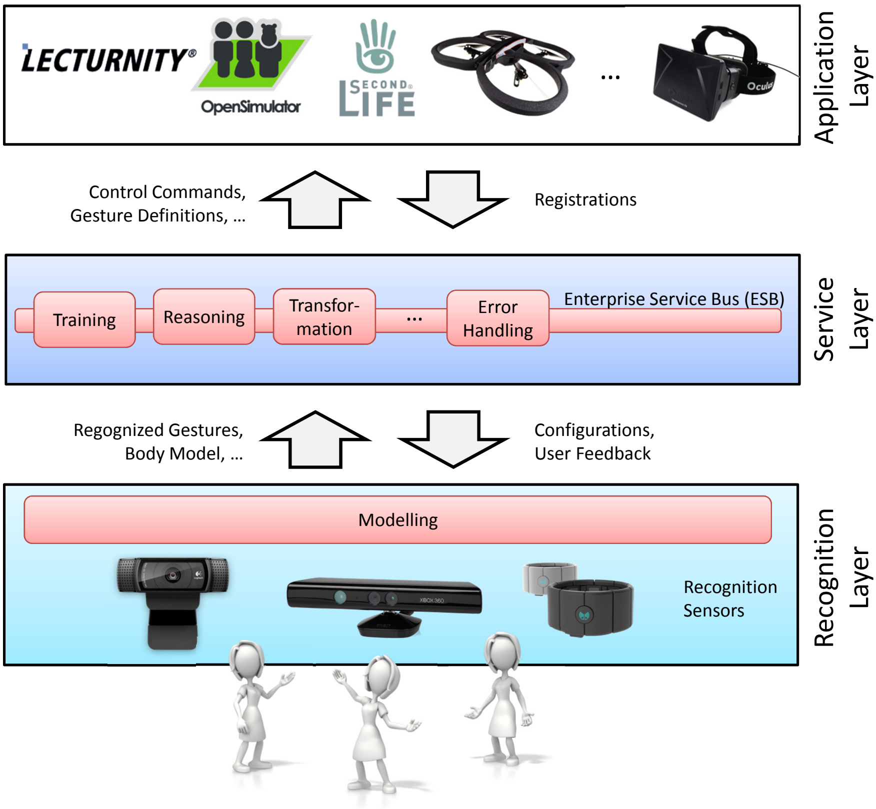



**Fig. 1.** The gesture control ecosystem separates applications and recognition system from each other by the use of a service layer

The bottom layer of the ecosystem manages all technology-specific mechanisms of gesture recognition. As today's gesture recognition tools typically come with own SDKs for common programming languages like Java or C#, the extensive handling of hardware issues is often already encapsulated by software APIs. Developers that integrate new technologies to the ecosystem usually work on the

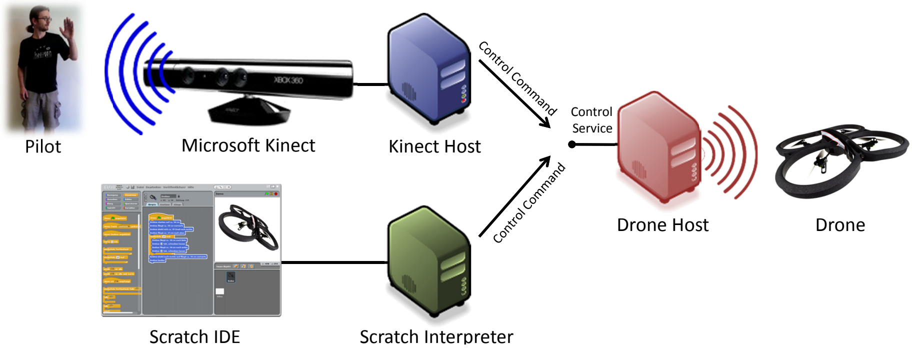



**Fig. 2.** The drone control case study uses a drone host as service provider from application layer

recognition layer. The service layer of the ecosystem systematically brings together recognition technology and applications. The basic integration mechanism on this layer is the use of an SOA that allows the encapsulation of implementation details behind functionality in terms of service interfaces. Beside the pure hosting of service endpoints to be accessed by the other layers, this layer should provide several basic services at a minimum. The application layer contains all hardware and software logic for specific use cases. Systems on this layer use only information and functionalities that are dedicated to their own functionality and that are provided by the service layer.

The overall ecosystem architecture allows a systematic combination of gesture recognition technologies and applications. The following case study is one example for a result of this systematic integration concept.

## 3 Case Study: Control of Physical Artefacts

In this case study, we aimed at the interpretation of a pre-defined gesture set as control commands for an AR.Drone 2.0 quadrocopter. The general architecture is shown in figure 2. On the recognition layer, the Kinect has been used. In addition, the integration of a MYO armband controller is currently in development. The application layer is represented by the quadrocopter and its drone host system. The drone host communicates with the drone using WiFi. As the AR.Drone 2.0 is not able to provide SOA services on its own, the drone host acts as provider of a set of REST services to control the quadrocopter.

Beside the gesture control itself, the ecosystem has also been used to control the drone application with another application: A Scratch-based IDE for childs has been integrated as an alternative controller for the AR.Drone 2.0. It uses the same services as the Kinect host and allows children to write simple programs (flight sequences) to be carried out by the drone.

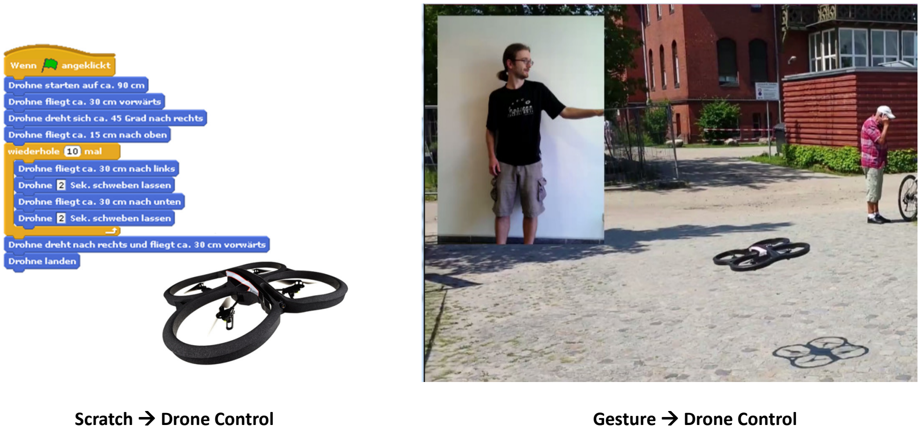



**Fig. 3.** The drone can be controlled by Scratch commands or live gestures of a human pilot

Figure 3 illustrates the results of this case study: The drone can be controlled by a Scratch program as well as a human gesturing pilot using the same infrastructure. Even in this semantic case study with realtime reactions as an important requirement, the ecosystem architecture performed well. Moreover, it showed its flexibility as the gesture control can easily be replaced by another input source like the scratch IDE.

## 4 Conclusion

This demo paper introduced an integrative NUI ecosystem, designed for but not limited to educational settings. The challenge of interoperability has been be addressed by a service layer that abstracts from technology-specific issues in terms of services of a Service Oriented Architecture. Thus, applications have been enabled to use gesture input without knowing the specific sensors. The ecosystem approach has been developed and evaluated with several case studies. One of them has been described in this paper: The control of an physical artefact by an NUI sensor as well as a graphical IDE used for teaching programming to children.

This work is one step towards an evolution of NUI experiments to a NUI ecosystem and a change from the current set of geek toys to a meaningful, multi-purpose toolkit for real educational scenarios.

## Reference

1. Norman, D.A.: Natural User Interfaces Are Not Natural. ACM Interactions 10, 6–10 (2010)

# Inquiring Rock Hunters

Maria Aristeidou, Eileen Scanlon, and Mike Sharples

Institute of Educational Technology,
The Open University, U.K.
{Maria.Aristeidou,Eileen.Scanlon,Mike.Sharples}@open.ac.uk

**Abstract.** The Inquiring Rock Hunters project is an example of Citizen Inquiry. Adult citizens designed and ran their own investigations in geology, through the online platform called nQuire, while collaborating with scientists. Analysis of the investigations shows patterns of collaboration and mentoring between novice, intermediate and expert geologists, however further work is needed to create a self-sustaining community of inquiry.

**Keywords:** Citizen Inquiry, Online Science Learning, Engagement with Science, Scientific Investigations, Community of Inquiry.

## 1 Overview

Citizen Inquiry for Inquiring Rock Hunters combines aspects from Citizen Science and Inquiry-Based Learning [1]. Important components of the design of the project are collaboration, knowledge sharing, peer review aligned to Citizen Science, as well as experimentation, discovery, critique and reflection associated with Inquiry-Based Learning. Citizen Inquiry addresses a need for citizens to adopt a sense of shared responsibility for scientific issues and become critical of the scientific information they receive [2].

The aim of Inquiring Rock Hunters is to support members of the public in designing and engaging in scientific investigations on rocks through an online community of inquiry. Geologists (experts and non-experts) having a shared interest in geology interact and exchange knowledge and methods supported and guided by the tools within the web-based inquiry environment, the nQuire platform [3][4].

Inquiring Rock Hunters is the first design study of a design-based research project which aims to explore the potential of Citizen Inquiry in engaging citizens with science and scientific investigation. Thus, the research question is shaped to this theme:

*"How can non-expert geologists engage in successful* ***Inquiry-Based Learning*** *through peer* **collaboration** *and* **mentoring** *by experts within* ***informal settings****?"*

For the data collection a survey and a System Usability Scale questionnaire were given to the participants, the published investigations were analyzed, and an online focus group discussion and online and face-to-face interviews were conducted.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 546–547, 2014.
© Springer International Publishing Switzerland 2014

## 2 Inquiring Rock Hunters

The project recruited 24 adult participants with interests in geology, of whom twelve created their own investigations. Seven investigations were complete (conclusions included) and three investigations received feedback without a follow-up discussion. The participants communicated through a forum and a chat. The usual pattern was for the amateurs to ask for help, the experts to guide and the intermediate geologists to be inactive. The participants reported that the platform was not very usable and that they needed a tutorial on how to operate it.

The results of the project confirm the active contribution of only some members; the creators of incomplete investigations do not seek for help within the community; there is no response to the feedback given on some investigations; the platform is abandoned once the investigation is completed. These issues emerged mostly because of the nQuire platform and the inquiry design which do not allow participants to conduct more than one investigation or collaborate with other members for their investigation. This leads to the possibility of people conducting individual investigations and quitting when they are finished or they are not able to finish.

## 3 Conclusions

The findings of Inquiring Rock Hunters project suggest that in the next phase of the study there is a need to improve the sense of community within the Citizen Inquiry members. A more intensive exploration of communities of practice and online communities may help build a self-sustaining community with a steady or increasing number of members and interactions.

## References

1. Aristeidou, M., Scanlon, E., Sharples, M.: A design-based study of Citizen Inquiry for geology. In: Katherine, M., Tomaž, K. (eds.) Proceeding of the Doctoral Consortium at the European Conference on Technology Enhanced Learning co-Located with the EC-TEL 2013 Conference, pp. 7–13. CEUR (2013)
2. McCallie, E., Bell, L., Lohwater, T., Falk, J.H., Lehr, J.L., Lewenstein, B.V., Needham, C., Wiehe, B.: Many Experts, Many Audiences: Public Engagement with Science and Informal Science Education. In: A CAISE Inquiry Group Report. Executive Summary. Center for Advancement of Informal Science Education (2009), http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED536432 (accessed June 15, 2009)
3. Mulholland, P., Anastopoulou, S., Collins, T., Feisst, M., Gaved, M., Kerawalla, L., Paxton, M., Scanlon, E., Sharples, M., Wright, M.: NQuire: Technological support for personal inquiry learning. IEEE Transactions on Learning Technologies 5(2), 157–169 (2012)
4. Scanlon, E., Anastopoulou, S., Kerawalla, L., Mulholland, P.: How technology resources can be used to represent personal inquiry and support students' understanding of it across contexts. Journal of Computer Assisted Learning 27(6), 516–529 (2011)

# Unsupervised Machine Learning Based on Recommendation of Pedagogical Resources

Brahim Batouche, Armelle Brun, and Anne Boyer

University of Lorraine,
Loria UMR 7503, France
{brahim.batouche,armelle.brun,anne.boyer}@loria.fr

**Abstract.** In E-learning context, we can recommend pedagogical resources to help learners. In this context, the recommender proposes the nearest resource(s) in term of similarity, but the scarcity of resources may affects seriously the quality of predictions. To make accurate predictions we begin in determining the scarce resources to be taken into account in the recommendation process. To achieve this objective we use the unsupervised neural network I2GNG (Improved Incremental Growing Neural Gas).

## 1 Introduction

The UOH (www.uoh.fr) portal is used as an interface with an e-learning repository, containing pedagogical resources. As everyone can freely navigate, it is important to assist the user by a recommender service.

We propose a recommender system, which relies on the last resources the user has consulted, and the description of the resource, if available. It is crucial to have the capability of making accurate predictions, because in the context of e-learning we cannot propose a service with a low quality of prediction, unlike commercial recommender service, such as amazon. The quality of prediction can be highly affected by the scarcity of resources, because it corresponds to a low coverage of resources, i.e., a bad diversity of resources. If it is an isolated resource (no similar resource exists), the proposed (or recommended) resource(s) can be dissimilar, and this is a problem.

## 2 Methodology

The task of the recommender system is similar to a librarian's task, which classifies all books by thematics. The user will be affected to the thematic which answers his request. Our recommender service classifies UOH resources in clusters (thematics) in order to give recommendation according to the followed thematic. If because of scarcity of resources in a specific thematic, the user has already see all resources (books), then the recommender (librarian) will have nothing to offer.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 548–549, 2014.
© Springer International Publishing Switzerland 2014

In this scenario, the librarian's knowledge allows to ensure the quality of prediction. This knowledge can be provided by an unsupervised machine learning I2GNG (Improved Incremental Growing Neural Gas) [1]. Wherefore, we use clustering not-only for its advantages observed by [2,3], but also to collect meta-data about a dataset of resources (knowledge about resources), e.g., classes of resources and links between classes. This knowledge presents a great advantage to solve the scarcity problem. It is the reason why we decide to provide recommendation according to the description of resources and also their meta-data.

We use a clustering method to determine scarce resources. We have to remind that scarce resources could be considered as noise for most of the clustering algorithms. Then we decide to use a dynamic clustering method allowing to create and remove clusters during the learning phase. Additionally, I2GNG allows to take into account, in real time, the change and update done in the dataset, without need to filter noises. The results of I2GNG shows distribution of resources and their scarcity, and the provided clusters are made up of a set of similar resources, whose can be used by the recommender. In summary, the use of I2GNG has fourfold objectives: (1) analyse scarcity of resources, (2) ensure the quality of prediction, (3) build knowledge-based recommender, (4) ensure scalable knowledge: neurones (or clusters) and connexions of neurones.

## 3 Results and Conclusion

The UOH dataset contains 1294 resources, which are treated by I2GNG. The results show that 36.78% of treated dataset are scarce i.e., isolated in their cluster. 63.22% of the rest are divided into 187 clusters, with an average size of 3.80 and a deviation of 3.24. 39 % of clusters may be merged in nearest future of use. The percentage of scarce resources (36.78%) corresponds to the probability of recommended resources with a low quality of prediction, and this without analysing the dataset, as our approach allows it[1].

## References

1. Hamza, H., Belaïd, Y., Belaïd, A., Chaudhuri, B.B.: Incremental classification of invoice documents. In: 19th International Conference on Pattern Recognition - ICPR, p. 4. IEEE, Tampa (2008) ISBN: 978-1-4244-2174-9
2. Kim, T.-H., Yang, S.-B.: Using attributes to improve prediction quality in collaborative filtering. In: Bauknecht, K., Bichler, M., Pröll, B. (eds.) EC-Web 2004. LNCS, vol. 3182, pp. 1–10. Springer, Heidelberg (2004)
3. Sarwar, B.M., Karypis, G., Konstan, J., Riedl, J.: Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering. In: Proceedings of the Fifth International Conference on Computer and Information Technology, vol. 1, Citeseer (2002)

[1] This work has been performed under the project Pericles (e-pericles.org).

# Toward Maximizing Access Knowledge in Learning: Adaptation of Interactions in a CoP Support System

Rachid Belmeskine[1,2], Dominique Leclet-Groux[1], and Ahlame Begdouri[2]

[1] MIS Laboratory (Modelization, Information, Systems),
University of Picardie, Amiens, France
rachid.belmeskine@hotmail.com, dominique.leclet@u-picardie.fr
[2] SIA Laboratory (Intelligent Systems and Applications), USMBA, Fez, Morocco
abegdouri@gmail.com

## 1 Problematic and Objectives

MAETIC [1] is a teaching method based on the realization of projects by the students and aims to allow them develop in parallel with know-how related to the target profession, professional skills as planning, negotiation, teamwork, communication, etc. The experiments [2] have shown that when a teacher deploys this method in his classroom, he may encounter difficulties due to the lack of expertise and experience. Then he may need help. The experiments have also shown that such aid could be provided by more experienced teachers in the use of the method. The teachers who use the MAETIC method therefore constitute a Community of Practice (CoP) whose the main objectives are: 1) the resolution of difficulties related to the use of MAETIC and 2) the exchange of practices related to this method. To Support this CoP, we interest on the design and development of an adaptive and mobile computer system that called NICOLAT (French acronym: iNformatIque COmmunautaire mobiLe et AdapTatif). Especially, NICOLAT aims to help these teachers members of The MAETIC CoP to resolve the difficulties that encounter during the use of MAETIC in classroom.

The MAETIC CoP is a community as well as the teachers often need solutions in real professional situations at their places of work (during the deployment of MAETIC in classroom). In similar conditions, the obstacles that influence negatively the rapidity and the continuity of interactions, aiming to resolve the difficulties encountered by the teachers, also influence their effectiveness and solutions utility. We postulate that these obstacles are related essentially to: 1) the problems of manipulation and familiarization with the new IT tools, 2) the response time to a request of difficulty resolution and 3) the access in case of mobility or lack of internet.

To limit these obstacles, we interest, via the adaptation of interactions, to 1) allow the CoP members to interact between themselves via popular tools used daily by them, 2) select only the members that could help to resolve a posted difficulty and solicit them in the tools used by them at the time of the publication, 3) develop adapted applications to the most used mobile devices and exploit the mobile telephony services in case of the lack of internet.

## 2 Structure of NICOLAT

NICOLAT consists of four layers which are:

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 550–551, 2014.
© Springer International Publishing Switzerland 2014

- ***The kernel:*** that incorporates the basic functionalities of a community environment personalized according to the specificities of the CoP MAETIC [3]. This kernel was developed based on the CMS (Content Management System) : Drupal.
- ***The RD (Resolving Difficulties) Layer:*** that allows to avoid similar helping requests by helping the teacher to solve his difficulties himself based on difficulties which have been solved in the past [3]. It is based on the Case Base Reasoning (CBR) and was developed as a Drupal module.
- ***The adaptation layers:*** In order to adapt the interactions between the CoP members, we designed two adaptation layers: ***1) The Adaptation of interactions layer*** and ***2) The Interface layer***.

## 3 Adaptation of Interactions in NICOLAT

The adaptation of interactions in NICOLAT aims to offer for each CoP MAETIC member the opportunity to use his favorite tools to interact with the system and the other members. Thus, it also aims, after the publication of a difficulty, to select only the CoP members who can help to find a solution and solicit them in the tools used by them at this time. To ensure this adaptation, we designed two layers: ***1) the Adaptation of Interactions (AI) layer*** which allows the management of interactions and the transformations of contents and ***2) the interface layer*** which provides interfaces, applications and software components to serve the AI layer and whose goal is to communicate with the CoP members. These two layers are considered as a bridge between the CCE and the tools used by the community members. The interaction tools that we chose to interconnect with the CCE are: Email, Facebook, SMS, Android and iOS devices. Among the components of the AI layer, we find the ***'Orchestration of the Adaptation'*** component which manages the adaptation by performing two complementary tasks: ***1) Task 1***: it selects the members that could help to resolve the difficulty posted based on the members profile information and the difficulty characteristics. ***2) Task 2***: it choices the tool used by each member selected and in which the system can solicit him based on his preferences (in terms of the interaction tools) and the history of his interactions.

## References


1. Talon, B., Leclet, D.: Dispositif pédagogique pour un apprentissage de savoir-faire: Dispositif pédagogique e-mallette MAETIC. Revue internationale des technologies en pédagogie universitaire (RITPU 2009) 5(2), 58–74 (2008)
2. Quenu-Joiron, C., Leclet, D.: How to Instrument a Community of Practice Dedicated to Project Based Pedagogy tutors: a Solution Based on Case Based Reasoning. In: 10th IEEE International Conference on Advanced Learning Technologies (ICALT 2010), Sousse, Tunisia, pp. 344–348 (2010)
3. Belmeskine, R., Begdouri, A., Leclet, D.: Resolution of difficulties approach for a Community of Practice members: design, implementation and experiment. International Journal of Research in Education Methodology 4(1), 409–422 (2013)

# An Assessment Methodology of Short and Open Answer Questions SOAQs

Safa Ben Salem, Lilia Cheniti-Belcadhi, and Rafik Braham

PRINCE Research Group, ISITCom, University of Sousse, Hammam Sousse, Tunisia
bensalem_safa@yahoo.fr, lilia.belcadhi@infcom.rnu.tn

**Abstract.** The short and open answer is a response written by a learner in natural language (until 50 words). We will propose a novel method of SOAQ assessment using concept maps and semantic similarity measure.

### Step 1: Analyzing the Text of the Answer

We use the Tree tagger [5] Part-Of-Speech tagger for annotating text with part-of-speech and lemma information. The output file of TreeTagger will be filtered by a program to remove the noisy words like the, an, a, or, two, etc. and to keep noun, adjective, adverb, verb and name as key words of the learner's answer.

### Step 2: Graphical Representation

As our ultimate objective is to assess SOAQs without losing the thread of idea nor the keywords used in the answer and also semantics, we chose to use the concepts maps as a knowledge representation which are called "Cmaps" [2]. In our case, the concept can be a noun, name, adjective or adverb or contain more than term in some case However, the link phrase is a verb. In order to extract more complex terms and nominal group in this step, we have to apply anaphoric resolution based on rule-based reasoning. The final concept map is composed of triplet " node-relation-node " which forms one significant "proposition", unit of meaning, sometimes called semantic unit (semantic links or links of meaning) [1].

### Step 3: Comparison between Concept Map Learner and Concept Map Model

This comparison process will be carried out with a semantic similarity measure in WordNet. Formally a hierarchical concept map [3] can be defined as a set:

$$CMap = \{C, R, T\} \quad (1)$$

Where: $C = \{c_1, c_2, c_3, \ldots c_n\}$ is a set of concepts. Each concept $c_i \in C$; $1 \leq i \leq n$ is a word or nominal group and it is unique in C. $R = \{r_1, r_2, r_3, \ldots r_m\}$ is a set of relationships among concepts. Each relationship $r_i \in R = (c_p, c_q)$; $p \neq q$; $1 \leq p \leq n$ ; $1 \leq q \leq n$ connects two concepts $c_p$ et $c_q$ $\in C$. $T = \{t_1, t_2, t_3, \ldots, t_n\}$ is a set of triplets. Each triplet $t_i \in T = (c_i, c_j, r_k)$; $i \neq j$; $1 \leq i \leq n$; $1 \leq j \leq n$; $1 \leq k \leq m$; $c_i$ et $c_j \in C$ and $r_k \in R$.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 552–553, 2014.
© Springer International Publishing Switzerland 2014

To compare learner's Cmap with model's Cmap we propose to compare the propositions as a first step. To understand this method clearly, let's consider that there are 3 propositions in learner's Cmap and 3 propositions in a model's Cmap. The initial step of the process can be shown with the help of the following figure.

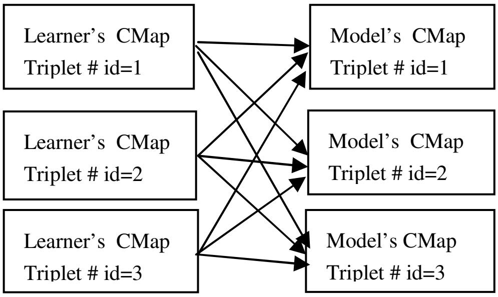



**Fig. 1.** Comparison between Triplets

We can't compute the similarity between triplets in bloc but we must divide it into terms. Based on the principle operating of WordNet similarity, it supported the same P-O-S of pair words so: the distance between a triplet from a learner's Cmap and the triplet from a model's Cmap is the sum of the distance between different terms of the triplet. Given T= ($c_{i1}$, $c_{i2}$, $r_k$) a triplet from learner's Cmap and T'= ($c'_{i1}$, $c'_{j2}$, $r'_k$) a triplet from model's Cmap, the distance Δ between them is defined as the sum of distances δ between all terms of triplet, like the following equation:

$$\Delta ( <C_{i1}, R_k, C_{i2}>, <C'_{j1}, R'_k, C'_{j2}> ) = \alpha\, \delta ( C_{i1},C'_{j1} ) + \beta\, \delta ( R_k,R'_k ) + \gamma\, \delta ( C_{i2},C'_{j2} ) \quad (2)$$

Where: $1 \leq i \leq n$; $1 \leq j \leq m$; with n: number of concepts of learner's CMap, m: number of concepts of model's CMap; $1 \leq k \leq p$; with p: number of relationship; ($\alpha$, $\beta$, $\gamma$) are constant.

We opted for semantic similarity measures based on the information content. This type is based on the assumption that the more common information two concepts and it is so effective. As soon as the task of comparison between triplets is completed, we will compute a score for the short and open answer question.

## References

1. Gouli, E., Gogoulou, A., Papanikolaou, K., Grigoriadou, M.: Compass: an adaptive web-based concept map assessment tool. Concept Maps: Theory, Methodology, Technology. In: Proc. of the First Int. Conference on Concept Mapping, Spain (2004)
2. Novak, J.D., Gowin, D.B.: Learning How To Learn. Cambridge University Press (1984)
3. Villalon, J.J., Calvo, R.A.: Concept map mining: A definition and a framework for its evaluation. Web Intelligence and Intelligent Agent Technology, 357–360 (2008)
4. Ben Salem, S., ChenitiBelcadhi, L., Braham, R.: Using Concept Maps in Knowledge Assessment of Short and Open Answer Questions. In: 6th International Workshop on Social and Personal Computing for Web-Supported Learning Communities SPeL (October 2013)
5. TreeTagger, `http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/`

# Estimating the Gap between Informal Descriptions and Formal Models of Artifact Flows in CSCL

Osmel Bordiés[1], Andreas Papasalouros[2], and Yannis Dimitriadis[1]

[1] Universidad de Valladolid, Valladolid 47011, Spain
obordies@gsic.uva.es, yannis@tel.uva.es
[2] University of the Aegean, Karlovassi 83200, Greece
andpapas@aegean.gr

**Abstract.** Teacher-designers incorporate constructs in their informal designs descriptions, which vaguely express the artifact flow dependencies as configuration constraints. This study analyzes through a corpus of CSCL designs, whether these constructs are supporting means for modeling, and the value of these elements is lost in the modeling process.

**Keywords:** Learning design, artifact flow, variability facets.

In advanced CSCL designs the definition of learning activity flows is complemented by incorporating the artifact flow, as a coordination mechanism, in order to satisfy the dependencies (e.g. temporal, resources) among actors or individual activities involved in a learning process [3]. However, modeling artifact flow for particular CSCL situations is an effort demanding task due to the intrinsic complexity of collaboration, and the lack of conceptual or technical support, especially when the number of groups and artifacts becomes high. Current reuse-based design approaches e.g. [4], promise to alleviate the cognitive load of teachers-designers by introducing abstract artifact flow models, which are the base upon which particular models are created or reused. Nevertheless, such approaches do not provide entirely satisfactory solutions to artifact flow in terms of reusability.

Constructs such as *"reports created by peers are reviewed"*, that provide the basis for the description and modeling of several specific configurations, are discarded through the abstract models formalization, and the particularization effort remains still high. We identify such constructs as *artifact flow variability facets* (from now on *variability facets*), whose unit is defined by an artifact flow situation where two or more activities are involved, and the relationships among them are expressed as configuration constraints or rules in terms of the coordination components: *Goals, Activities, Actors* and *Dependencies* [2]. The formal representation of the *variability facets* may help teachers to complete CSCL scenarios with artifact flow through the provision of technical and conceptual support. However, the limited expressiveness of modeling languages in use, such as IMS LD or BPEL, impedes the formalization of such constructs.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 554–555, 2014.
© Springer International Publishing Switzerland 2014

We performed the study by analyzing 30 CSCL scenarios descriptions extracted from several professional development workshops that took place in 2012 and 2013 at the University of Valladolid, Spain, where higher-education teachers authored realistic designs inspired by their own courses. The teachers were provided with a few design routines and well known collaboration patterns as starting points in order to encourage them to pose their own teaching-learning scenarios. The study follows a mixed analysis approach and starts with the identification of *variability facets* based on the pattern phases upon which specific coordination components are defined. For each pair of activities linked by an artifact flow dependency, we have identified whether they have been described through *variability facets*, and whether they are completely described in terms of concrete components settings or not. We detected 230 learning flow phases, out of which 150 are involved in *variability facets* constructs. The 45% of the *facets* occurrences are completely defined in abstract terms while the rest incorporates particular group settings. It is remarkable that 81% of the artifact flow situations are not particularized through the descriptions, neither formalized in abstract models using current modeling specifications [3], and they should be lately configured and deployed by the teachers. Furthermore, the specification of constraints for flexible workflow is still an open issue [1].

In this study we have illustrated the gaps that exist between the informal descriptions of artifact flows, the corresponding complete descriptions, as well as their formal representations. The *variability facets*, which are vaguely expressed as sets of high level customization constraints, were found to be present in most of the designs represented in the corpus. Also, the analysis showed that they have a high potential towards an effective particularization of the CSCL models.

**Acknowledgments** This research has been partially funded by the Autonomous Government of Castilla and León, Spain (EDU/346/2013), the Spanish Ministry of Economy and Competitiveness (TIN2011-28308-C03-02) and the European Education, Audiovisual and Culture Executive Agency (531262-LLP-2012-ES-KA3-KA3MP).

## References

1. Bajwa, I., Lee, M., Bordbar, B.: SBVR Business Rules Generation from Natural Language Specification. In: AAAI Spring Symposium Series (2011)
2. Malone, T.W., Crowston, K.: What is Coordination Theory and How Can It Help Design Cooperative Work Systems? In: Proceedings of the 1990 ACM Conference on Computer Supported Cooperative Work, CSCW 1990, pp. 357–370 (1990)
3. Miao, Y., Burgos, D., Griffiths, D., Koper, R.: Representation of Coordination Mechanisms in IMS Learning Design to Support Group-based Learning. In: Lockyer, L., Bennet, S., Agostinho, S., Harper, B. (eds.) Handbook of Research on Learning Design and Learning Objects: Issues, Applications and Technologies, pp. 330–351. IDEA Group (2008)
4. Palomino-Ramírez, L., Bote-Lorenzo, M.L., Asensio-Pérez, J.I., Vignollet, L., Dimitriadis, Y.A.: LeadFlow4LD: A Method for the Computational Representation of the Learning Flow and Data Flow in Collaborative Learning. Journal of Universal Computer Science 19(6), 805–830 (2013)

# An Elicitation Process for Know-How Transfer of Teachers' Usual Practices by a Task/Method Approach

Thierry Condamines[1], Dominique Leclet[1], and Hanane Barramou[2]

[1] MIS Laboratory,
University of Picardie Jules Verne, Amiens, France
{thierry.condamines,dominique.leclet}@u-picardie.fr
[2] LSIA Laboratory, Faculty of Science and Technology, Fez, Morocco
hanane.barramou@gmail.com

**Abstract.** Knowledge exchange and capitalization in a Community of Teachers is not easy to perform. Online forums have shown their limits and the need for systems dedicated to Communities of Practice is real. In this paper, we propose a bottom-up approach for a platform model to allow practice exchange in a teachers' Community of Practice following a Task/Method paradigm for practice description and a community driven elicitation process.

**Keywords:** Knowledge sharing, Knowledge Elicitation, Task/Method Paradigm, Teacher Training, Communities of Practice.

## 1 Introduction

Our work aims to provide a teachers' community with a collaborative platform [1], which will support knowledge elicitation about teaching practice (driven by interactions between peers, and capitalized into individual memories with the associated context) and sociability and member participation in a Web 2.0 approach.

## 2 Task/Method Paradigm and Community Driven Elicitation

Task/Method paradigm and Hierarchical Task Model are well known in AI [3] and learning environment design [4]. They can be used to represent a reasoning model: how a task (what the teacher have to do) can be performed, that is the method describing one way to perform the task. The decomposition of a task can be modeled as a hierarchical tree as shown in Fig. 1. But the conversion process from tacit to explicit knowledge is a social process and many techniques in knowledge elicitation are based on interviews of experts by a knowledge engineer. One of these techniques is the "Why How Technique" [4] we used to allow members of the community to select predefined questions about a task. For such a scenario we used a community management based on a user profile, to put people with the same concern into contact.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 556–557, 2014.
© Springer International Publishing Switzerland 2014

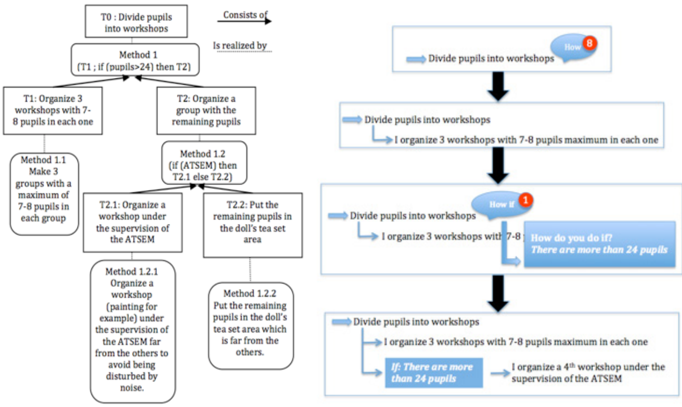



**Fig. 1.** Task/Method decomposition of a teacher's practice and community driven elicitation

## 3 Conclusion

In this paper, we have proposed a Task/Method approach with a community driven elicitation process for practice exchange and capitalization in a Community of Practice. By connecting teachers with similar professional context we induce a progressive refinement of practice description. And by broadcasting in the community a rich variety of practices, we hope to favor innovation.

**Acknowledgement.** This work was supported by Region of Picardy, France.

## References

1. Condamines, T.: A Collaborative Platform Supporting Knowledge Sharing and Problem Solving for Teacher Professional Development. In: 9th Int. Conf. on Computer-Supported Collaborative Learning (CSCL 2011), Hong-Kong, July 4-8, vol. 1, pp. 486–493 (2011)
2. Trichet, F., Tchounikine, P.: DSTM: a Framework to Operationalize and Refine a Problem-Solving Method modeled in terms of Tasks and Methods. International Journal of Expert Systems With Applications (ESWA) 16, 105–120 (1999)
3. El-Kechaï, N., Després, C.: A Plan Recognition Process, Based on a Task Model, for Detecting Learner's Erroneous Actions. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) ITS 2006. LNCS, vol. 4053, pp. 329–338. Springer, Heidelberg (2006)
4. Graesser, A.C.: How to catch a fish: The memory and representation of common procedures. Discourse Processes 1, 79–89 (1978)

# EAGLE – Open Data and Linked Data Architecture of an Enhanced Government Learning Platform*

Violeta Damjanovic[1], Dietmar Glachs[1], Nikolay Tcholtchev[2], Eric Ras[3], and Eric Tobias[3]

[1] Salzburg Research, Austria
[2] Fraunhofer Institute for Open Communication Systems, Germany
[3] Public Research Centre Henri Tudor, Luxembourg
{violeta.damjanovic,dietmar.glachs}@salzburgresearch.at, nikolay.tcholtchev@fokus.fraunhofer.de, {eric.ras,eric.tobias}@tudor.lu

**Abstract.** We present the architecture of an Open Data and Linked Data platform for "EnhAnced Government LEarning" (EAGLE), which aim is to help local governments to keep up with fast-changing trends in public administration (PA) by adopting technology-enhanced learning (TEL) methods. EAGLE uses Linked Open Data tools: Apache Marmotta and Apache Stanbol.

**Keywords:** Linked Open Data, semantic enhancement, government learning.

## 1 Introduction

In our pre-study on learning in public administration (PA), in rural local government, conducted in Luxembourg, Montenegro, Germany, Austria, and Ireland in 2012/13, the following problems have been identified: "*Lack of timely access to learning content, lack of well-established learning processes at the workplace, lack of change management procedures […]*" followed by the fact that " *[…] in 2012, 86 regulations and 67 laws were changed in Germany; 54 regulations and 33 laws were already changed in 2013.*" Hence, our motivation is to create a learning platform, which eases adoption of changing trends in PA. Herein, we present the architecture of the EAGLE Open Data and Linked Data (LD) learning platform, which builds on top of existing open source tools and frameworks, and develops a novel search, presentation and navigation strategy for education in PA.

## 2 EAGLE Architecture

The EAGLE learning platform architecture follows an open source policy, and reuses existing open source tools and components, such as relevant public sector datasets developed in LATC and LOD2 projects, as well as LD generated in different PA

* The research leading to these results has received funding from the EC FP7 (FP7/2007-2013) under grant agreement N° 619347.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 558–559, 2014.
© Springer International Publishing Switzerland 2014

context. It builds on top of several existing tools for the creation, linking, and publication of LD, and integrates open source tools: Apache Stanbol, which allows for semantic analysis, and Apache Marmotta, a LD platform [1]. The aim of the EAGLE architecture is to support aggregation, refinement, and usage of various Open Educational Resources (OER). The architecture specifies a set of components (see Figure 1): (i) *OER Linked Data,* meant for information management, (ii) *OER Data Registry,* for data management, based on the OpenData CKAN standard, and (iii) *Automatic Item Generation (AIG)*, for self-assessment of information literacy in PA.

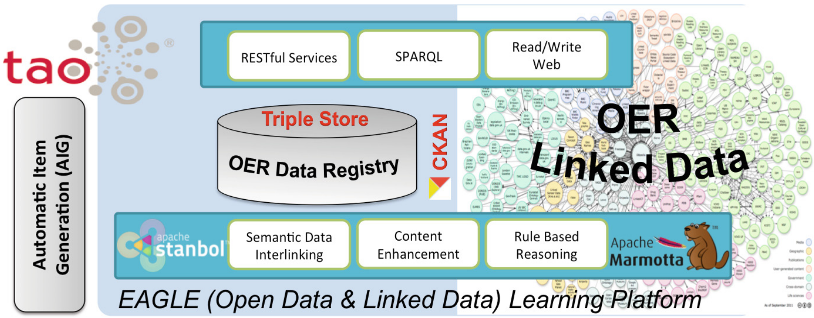



**Fig. 1.** High-level architecture of the EAGLE Open Data & Linked Data learning platform

To realise OERs as LD, we apply the following steps: (i) *Data modelling:* we develop an OER ontology to be used in PA; (ii) *Data reusing:* we either import structured data from the existing LD sources, or transform non-structured sources first into RDF, then align structured RDF data to the OER ontology; (iii) *Data interlinking:* we perform semantic lifting and content enhancement of the data to be stored, or modified; (iv) *Data storage:* all content enhancements and data aligned to the OER ontology are stored in the OER Data Registry; and (v) *Data publication:* the data from the EAGLE platform are available to EAGLE (as well as LD-) services, either via SPARQL or via RESTful services.

## 3 Conclusions and Future Work

Incorporating LD and Semantic Web technologies for learning in PA offers several benefits, such as collaborative creation and sharing of knowledge, improved interactivity due to concept-based navigation and search capabilities. In EAGLE, the major technical challenge will be to provide a smooth integration of the existing open source platforms/tools with PA services. One of the first steps is the enhancement of the OER metadata scheme to support CKAN for the *OER Data Registry* component.

## References

1. Kurz, T., Güntner, G., Damjanovic, V., Schaffert, S., Fernandez, M.: Semantic Enhancement for Media Asset Management Systems: Integrating the Red Bull Content Pool into the Web of Data. In: Multimedia Tools and Applications, MTAP (2012)

# Towards the Identification and Formalization of LMS Instructional Design Languages

Nour El Mawas, Lahcen Oubahssi, Pierre Laforcade, and Aymen Abedmouleh

Laboratoire d'Informatique de l'Université du Maine (LIUM), EA 4023
LUNAM Université, Université du Maine
72085 Le Mans, France
Surname.Name@univ-lemans.fr

**Abstract.** In recent years, researchers have begun to formalize LMSs instructional languages in order to specify models in conformance with their underlying infrastructure. To our knowledge, there is no proposition that focuses on identifying an explicit process to formalize such languages. The research presented in this paper aims to define the necessary analysis and steps for the identification and formalization of an LMS instructional design language. Academic institutes have adopted LMSs; however, many teachers have difficulty using LMSs to create learning designs. Teachers are not familiar with implicit learning design domain of LMSs. They need solutions to bridge the gap between their educational intentions and the pedagogical scenarios proposed by the LMS at their disposal. In our work, we aim at supporting practitioners to overcome these LMSs' obstacles in order to help them in focusing on the design of learning situations. Our contribution consists in extracting, identifying, and formalizing the LMS implicit instructional design language by proposing a meta-model to the design of learning situations.

We propose a process to identify and formalize the instructional language of LMSs. Our approach is based on four viewpoints: the macro-HMI analysis, the factorization of HMI-macro model, the functional analysis and the micro analysis.

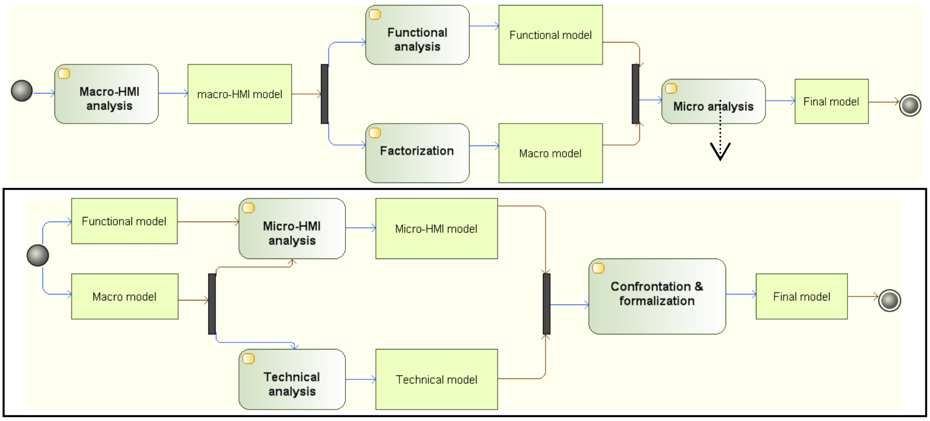



**Fig. 1.** Analysis process of the instructional design language

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 560–561, 2014.
© Springer International Publishing Switzerland 2014

The first viewpoint consists of HMIs analysis according to two strategies: (1) the analysis of existing situations on the platform and (2) the analysis of interfaces related to the specification of new situations. After the macro-HMI analysis, we factorized the macro-HMI model in order to obtain the simplified macro model. The functional analysis focuses on the identification of LMS existing functions. The last viewpoint concerns the micro analysis of the LMS instructional design language. It is composed of three steps: the micro-HMI analysis, the technical analysis, and the confrontation of micro-HMI and technical models.

We apply our process on the Moodle 2.4 platform. The application of **macro-HMI analysis on Moodle** consists of identifying interfaces related to course design. We analyzed interfaces titles and navigation paths / URLs. We browsed all the links in a specific interface. After the macro-HMI analysis, we applied the **factorization process**. For example, we noticed that all activities/resources had the common attributes: "*commonModuleSettings*", "*restrictAccess*", and "*activityCompletion*". So we moved these attributes to the Activity/Resource class. Based on the macro-HMI model, we proceeded to the **functional analysis on Moodle**. We divided each interface to several areas. Then, for each area, we studied the graphical interface components to identify functionalities related to instructional design. For example, from the main interface of a Moodle course, a teacher can show/hide/move a section. He can modify the course description, and manage different groups. He can also add an activity/resource in a specific section… The application of **IHM-micro analysis** is about characteristics identification of instructional design elements. It is based on the macro and functional models. The **technical analysis** consists in analyzing the Moodle database. Our goal is to identify the Moodle instruction design language from a technical viewpoint to approve the relevant of specific data for this language. This analysis consists in specifying the reduced Conceptual Data Model for Moodle in relation with the instruction design. We have reviewed all Moodle database tables. The micro-HMI analysis and the technical analysis have specified two Moodle instructional design models according to two different viewpoints. In this step, we are interested in the **confrontation of these models** to formalize the Moodle instructional design language. We think that the use of only one analysis method presents many negative points. Indeed, the micro-HMI model depends directly on the Moodle analyst competence. This means the possibility lack of pedagogical attributes. Similarly, the technical analysis is not an easy task. Many data structures are not explicitly reported when creating the database.

In conclusion, we present in this paper a meta-model-based approach for identifying and formalizing LMS languages. We apply our proposed approach on the Moodle platform. The meta-model will be used as a basis for the development of the external editor. It will guide and generate most of the final code for the editor. This will facilitate the use of LMSs and allow teachers becoming more familiar with e-learning.

**Acknowledgement.** This work and submission are funded by the French GraphiT project [ANR-2011-SI02-011] (http://www-lium.univ-lemans.fr/~laforcad/graphit/).

# Taking on Different Roles: How Educators Position Themselves in MOOCs

Rebecca Ferguson and Denise Whitelock

IET, The Open University, Walton Hall, MK7 6AA, UK
{firstname.lastname}@open.ac.uk

**Abstract.** Educators in massive open online courses (MOOCs) face the challenge of interacting with tens of thousands of students, many of whom are new to online learning. This study investigates the different ways in which lead educators position themselves within MOOCs, and the various roles that they adopt in their messages to learners. Email messages from educators were collected from six courses on FutureLearn, a UK-based MOOC platform that had 26 university partners at the time. Educator stance in these emails was coded thematically, sentence by sentence. The resulting typology draws attention to the different ways in which educators align themselves in these settings, including outlining the trajectory of the course, acting as both host and instructor, sometimes as fellow learner, and often as an emotionally engaged enthusiast. This typology can be used to explore relationships between educator stance and variables such as learner engagement, learner test results and learner retention.

MOOCs were derived from connectivist theory, which emphasises the role of social context within learning [1]. The originators of connectivist MOOCs emphasised that 'you are teaching while you are learning' [2]. However, most MOOCs are not connectivist, so what is the educators' role when they are talking to and guiding thousands in an unfamiliar setting? 'Qualitatively different approaches to teaching are associated with qualitatively different approaches to learning' [3], so we would expect to see some changes in these new environments.

Goffman was clear about the role of presentation of self in any interaction: 'Information about the individual helps to define the situation, enabling others to know in advance what he will expect of them and what they may expect of him' [4, p1].

We conducted a preliminary study to investigate educator stance in six MOOCs presented on the FutureLearn platform and to understand how their communications can encourage participation and promote confidence. We took 'stance' to mean the alignment of the educator to learners/educators/content, signaled by linguistic cues or by changes to the frame of reference.

Our focus was on 'lead educators', the university-based individuals who signed off emails associated with the course. These are the educators most likely to be encountered by learners on FutureLearn, because these emails are sent weekly, or more frequently, to everybody registered on the course.

The study required a subtle yet rich instrument to encompass the lead educators' alignment to a very large group of students. Previous categorisation systems, which

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 562–563, 2014.
© Springer International Publishing Switzerland 2014

included those used for the study of group interaction sessions [5], tutor comments [6] and online tutor assistance [7] proved to have limitations in this context. We therefore coded the emails thematically, sentence by sentence, resulting in the category system outlined here, which we then used to consider educator stance in different MOOCs.

- Academic grouping: member of a grouping within the university
- Assessor: assessing student work
- Course team member
- Emotionally engaged
- Evaluator: evaluating the course
- Explainer: justifying why the course is structured as it is
- Group: member of the group of learners on this course
- Host, or part of a team of hosts
- Instructor: providing instructions and / or options
- Lead educator, on this course
- Outliner: outlining what is happening on the course
- Recommender: recommending resources, including course materials and URLs
- Social media user
- University member

This typology draws attention to the different ways in which educators align themselves in these settings, including outlining the trajectory of the course, acting as both host and instructor, sometimes as fellow learner, and often as an emotionally engaged enthusiast. The typology also identifies variation in educator stance between courses. Some were concerned with building a sense of community and obviously considered themselves to be part of that community, while others remained more aloof.

This preliminary research offers a new lens with which to explore teaching and learning activity within MOOCs.

## References

1. Siemens, G.: Connectivism: A Learning Theory for the Digital Age. elearnspace (2004), http://www.elearnspace.org/Articles/connectivism.htm
2. McAuley, A., Stewart, B., Siemens, G., Cormier, D.: The MOOC Model for Digital Practice (2010), http://davecormier.com/edblog/wp-content/uploads/MOOC_Final.pdf
3. Trigwell, K., Prosser, M., Waterhouse, F.: Relations between Teachers' Approaches to Teaching and Students' Approaches to Learning. Higher Education 37, 57–70 (1999)
4. Goffman, E.: The Presentation of Self in Everyday Life. Allen Lane, London (1969)
5. Bales, R.F.: A Set of Categories for the Analysis of Small Group Interaction. American Sociological Review 15, 257–263 (1950)
6. Whitelock, D., Watt, S.: Open Mentor: Supporting Tutors with Their Feedback to Students. Paper Presented at the 11th CAA International Computer Assisted Assessment Conference, Loughborough, UK, July 10/11 (2007)
7. Angeli, C., Valanides, N., Bonk, C.J.: Communication in a Web-based Conferencing System: The Quality of Computer-Mediated Interactions. British Journal of Eduational Technology 34(1), 31–43 (2003)

# Potential of EPUB3 for Digital Textbooks in Higher Education

Christian Gailer[1], Martin Ebner[1], Michael Kopp[2], Elke Lackner[2], Michael Raunig[2], and Alexei Scerbakov[1]

[1] Institute for Information Systems and Computer Media, Graz University of Technology, Inffeldgasse 16c, 8010 Graz, Austria
{christian.gailer,a.scerbakov}@student.tugraz.at, martin.ebner@tugraz.at
[2] Academy for New Media and Knowledge Transfer, University of Graz, Liebiggasse 9, 8010 Graz, Austria
{michael.kopp,elke.lackner,michael.raunig}@uni-graz.at

**Abstract.** The e-book market is currently in a strong upswing. This research study deals with the question which practical uses the e-book format EPUB3 offers for (higher) education. By means of a didactic content analysis, a range of interactive exercise types were developed as a result of conversations with teachers. For this purpose, a didactic and technical concept has been developed. Different kinds of exercises were prototypically implemented in an e-book. In summary, it can be remarked that EPUB3 is suitable for a variety of different exercises and that it is able to serve as a basic format for forthcoming digital textbooks. Furthermore the openness of EPUB3 will assist Open Learning and Teaching in a meaningful way.

**Keywords:** epub, e-book, interactive, enhanced, e-learning.

## 1 Introduction

In our publication we are strongly interested in the open standard format EPUB (electronic publication). Due to the fact that Open Learning and Teaching are strongly based upon Open Educational Resources (OER) our research concentrates on open standards such as EPUB3. The International Digital Publishing Forum (IDPF) published the EPUB3 standard in 2011 [1]. This new format opens up numerous opportunities for teaching by utilizing state-of-the-art technology.

This research addresses the question of what content is suitable in the field of (higher) education [3] and how the didactic and technical creation process should look like. For this purpose, a field study is carried out and a prototype of a French textbook is developed. Thereby, common work sheets and exercises have to be implemented into the e-book to point out the strengths and weaknesses of the format in a technical and a content-related sense.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 564–565, 2014.
© Springer International Publishing Switzerland 2014

## 2 The EPUB3 Prototype

Talking to language teachers, we conducted a didactical content analysis of the teaching contents, which should take into consideration the different learning styles and types, the different levels of knowledge and the four communicative language competences stated by the Common European Framework of Reference for Languages (CEFR). The exercises should not be only drill and practice exercises in a behavioristic way [2], but should be varied, appealing, interactive, motivating and multimedia-based. Last but not least, the exercises should be auto corrective in order to assure immediate feedback.

JavaScript was employed as the programming language to implement interactive exercises and contributes to interactivity in e-books. In addition, it is of high importance for both creating and changing content. The core of the system is patterned according to the Model-View-Controller concept.

Considering the didactical concept as well as technical possibilities different exercises have been created. These can be divided into four main categories: *Drag&Drop*, *Text Assignments*, *Quizzes* and *Selection Tasks*.

After presenting the prototype to a wider audience (of language learners, teachers and non-language learners) they were at first sight impressed by the high level of interactivity and multimedia implementation. At the second sight the teachers were delighted by the appealing layout and the fact that the exercises covered the different language competences and proficiency levels as stated by the CEFR.

## 3 Conclusions

This research shows that a wide range of content can be implemented in EPUB3 format. It can be pointed out that the strength is a high level of interactivity with the user. This results in a high level of motivation, which could be demonstrated by implementing a prototype. Since EPUB3 is an open format, this has a positive impact on its spread. Even now it is supported by platforms and devices from different manufacturers, although not yet fully. According to the current state, extensive knowledge in the areas of HTML5, CSS3 and JavaScript, as well as EPUB specifications are necessary for the creation of EPUB3 publications.

## References

1. Conboy, G., Garrish, M., Gylling, M., McCoy, W., Makoto, M., Weck, D. (eds.): EPUB 3 Overview. IDPF Recommended Specification (October 11, 2011), `http://www.idpf.org/epub/30/spec/epub30-overview.html` (accessed February 2014)
2. Morgret, K., Weber, N., Woo, L.: A Behaviorist Approach To Spelling. Learning (2001)
3. Nagler, W., Ebner, M., Scerbakov, N.: Reading and Learning with Any Device University Content for E-Books and E-Readers. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications, pp. 1775–1782 (2011)

# Examat: Intelligent Exercise Services for Mathematics Learning

Giorgi Goguadze

Leuphana University, Lüneburg, Germany
goguadze@leuphana.de

**Abstract.** One of the major achievements of Intelligent Tutoring Systems is the ability to provide deep diagnosis and intelligent feedback upon user actions in an inner loop of an interactive exercise. In this work, we develop a platform that hosts remote interactive exercises that can be plugged into any browser-based environment.

The system supports two feedback modes - a training mode with step-by-step evaluation of student's answer with immediate error-feedback and hint sequences, and a homework or assessment mode, in which a student can enter a complete multi-step solution of a problem, submit it into the system and receive delayed feedback to each made step. Such a detailed analysis of student solutions is provided by the context-aware domain modules used as back-engines.

## 1 Introduction

Since the raise of Intelligent Tutoring Systems the most important part of each ITS is an exercise module. Interactive exercises with intelligent feedback train learners to improve their procedural knowledge contributing to the key factor of knowledge acquisition according to ACT-R theory [1]. Several decades since the creation of ITS technology many systems are widely used in US and some are used in Europe as well, but this usage is still restricted due to several reasons. First of all, it is difficult for the lecturers to install and maintain such highly complex systems and universities do not provide enough support for this task. For example, an intelligent learning environment Math-Bridge[1] has been successfully tested in several universities in Europe, but its usage has been a subject to constant maintenance and content debugging, which in the long term led to abandoning the system as the technical support staff was not available any more. Secondly, the lecturers who would like to use interactive exercises usually do not need to use the complete ITS, but rather embed the exercises in their own lecture scripts or LMS systems. It is in general not an easy task to provide single exercises for the usage outside of an ITS environment. Existing systems such as STACK or MathDox[2] offer Web-based interactive exercises using semantic evaluation by a remotely accessed Computer Algebra System, but can only provide a shallow correct/incorrect feedback.

[1] `http://www.math-bridge.org`

[2] See`http://stack.bham.ac.ukandhttp://www.mathdox.orgresp`

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 566–567, 2014.
© Springer International Publishing Switzerland 2014

We address these limitations and offer an interactive exercise module Examat, providing intelligent interactive exercises for various topics in introductory university mathematics, which can be incorporated into any web-based content.

This work is based on the previous experience of the author with the development of an intelligent learning environment Math-Bridge, which remotely connects to external educational services (we use IDEAS domain reasoners[3]) providing intelligent diagnosis upon student actions [2].

## 2 Exercises with Two Modes of Feedback

We offer two modes of interaction with Examat exercises - the training mode, and the homework/assessment mode.

In the training mode the students receive correctness and error feedback at each step as well as hints and worked solutions of problems on their own solution path. The next step hints and worked solutions are generated using the application of expert rules of the external rule engine to the current step, and the error feedback is generated from the so-called buggy-rules that the external rule engine matches against the erroneous user input.

In a homework or assessment mode the students can enter a complete multi-step solution to a problem into the system, submit it and immediately receive an evaluation of the complete solution in the same way as it is usually done by the teacher, who evaluates not only the final result, but also intermediate steps. The error feedback is given for each incorrect step and the rest of the solution is further corrected given the current erroneous step was correct.

The system possesses an easy to use authoring tool, which consists of a graphical template, in which the authors have to fill in the metadata of an exercise, and its task, and the rest is generated automatically.

In the following months we expect to achieve the integration of the tool into the commonly used LMS platforms, as well as develop a collection of exercises for bridging-courses content to be integrated into a German online platform VEMINT[4] for remedial mathematics.

By providing intelligent and easy to use online exercises we expect to increase the impact of ITS in the community of mathematics educators in Europe.

## References

1. Corbett, A.T., Koedinger, K.R., Anderson, J.R.: Intelligent tutoring systems. In: Helander, M., Landauer, T.K., Prabhu, P.V. (eds.) Handbook of Human-Computer Interaction, 2nd edn., ch. 37, pp. 849–874. Elsevier Science B.V., Amsterdam (1997)
2. Goguadze, G.: Semantic evaluation services for web-based exercises. In: Spaniol, M., Li, Q., Klamma, R., Lau, R.W.H. (eds.) ICWL 2009. LNCS, vol. 5686, pp. 172–181. Springer, Heidelberg (2009)

[3] See http://ideas.cs.uu.nl

[4] See http://www.vemint.de

# Employing Speech to Contribute to Modelling and Adapting to Students' Affective States


Beate Grawemeyer[1,⋆], Manolis Mavrikis[2], Alice Hansen[2], Claudia Mazziotti[3], and Sergio Gutierrez-Santos[1]

[1] London Knowledge Lab, Birkbeck College, University of London, UK
{beate,sergut}@dcs.bbk.ac.uk
[2] London Knowledge Lab, Institute of Education, University of London, UK
{m.mavrikis,a.hansen}@ioe.ac.uk
[3] Institute of Educational Research, Ruhr-Universität Bochum, Bochum, Germany
{claudia.mazziotti}@rub.de



**Abstract.** Affect plays a significant role in students' learning behaviour. Positive affective states can enhance learning, while negative ones can inhibit it. This paper describes how we provide intelligent support in a learning platform based on affect states. We discuss two components: an affective state detector to perceive affective states in speech during interaction with the platform; and an affective state reasoner to provide support, which aims at aligning the learner's personal goal with the learning task to evoke positive affective states for an enhanced learning experience.


## 1 Introduction

Our aim is to build a learning platform for elementary education that integrates speech recognition for children in order to enable natural communication. We report on the development of an affective state detector capable of inferring affective states in children's speech. This is then used by the affective state reasoner to provide intelligent support that is responsive to the affective state.

As described in [1] affective states interact with and influence the learning process. While positive affective states contribute towards constructive learning, negative ones can lead to challenges. The learning process includes a range and combination of positive and negative affective states. This work contributes towards the understanding of how intelligent support can be used to turn negative affective states into positive ones.

## 2 Modelling and Adapting to Affective States

We adopt the layered approach designed by [4] for the development of the intelligent support. The input for the analysis layer is the result of the speech


⋆ This research has been funded by the EU in FP7 in the iTalk2Learn project (318051). Thanks to all our ITalk2Learn colleagues for their support and ideas.



C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 568–569, 2014.


© Springer International Publishing Switzerland 2014

recognition software by Sail Labs [2]. This software is trained specifically with children's voices. It creates an array of words based on speech input. This word array is used by the affective state detector for classification. The classification is based on the 'Bags-of-Words' model (e.g. [3]). We apply a naive Bayes classifier to classify the affective states. The system then passes the result of the classification to the reasoning layer, which uses this input to make decisions on how adaptive feedback should be provided, via the affective state reasoner.

As described in [5] positive emotions or affective states facilitate students' self-regulation of learning, figuring in meta-cognitive and meta-emotional strategies for adapting learning to goals and task demands. In contrast, negative emotions can lead to reliance on external guidance. A match between personal goals and learning tasks can produce positive affective states, while a mismatch can produce negative states [6]. Our affective state reasoner includes rules about how negative affective states can be transformed into positive ones by aligning the student's reasoning process with the learning task. The rules are based on Wizard of Oz studies [7] where the platform was used as a tool to investigate what type of support is effective for a particular affective state.

## 3 Conclusion and Future Work

We have developed a system that is able to provide adaptive feedback based on affective state. The next stage in our research will be to evaluate the system with respect to students' affective states. This will include an assessment of the effectiveness of the affective state detector and the affective state reasoner with regard to students' learning experience and performance.

## References

1. Porayska-Pomsta, K., Mavrikis, M., D'Mello, S.K., Conati, C., de Baker, R.S.J.: Knowledge Elicitation Methods for Affect Modelling in Education. I. J. Artificial Intelligence in Education 22, 107–140 (2013)
2. Sail Labs, `http://www.sail-labs.com`
3. Schuller, B., Müller, R., Lang, M., Rigoll, G.: Speaker independent emotion recognition by early fusion of acoustic and linguistic features within ensemble. In: Proc. Interspeech, Lisbon, Portugal, pp. 805–808 (2005)
4. Gutierrez-Santos, S., Mavrikis, M., Magoulas, G.D.: A Separation of Concerns for Engineering Intelligent Support for Exploratory Learning Environments. Journal of Research and Practice in Information Technology 44(3), 347–360 (2012)
5. Pekrun, R.: The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. J. Edu. Psych. Rev., 315–341 (2006)
6. Ahmed, W., van der Werf, G., Kuyper, H., Minnaert, A.: Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. Journal of Educational Psychology, 150–161 (2013)
7. Grawemeyer, B., Mavrikis, M., Gutierrez-Santos, S., Hansen, A.: Interventions during student multimodal learning activities: which, and why? In: Proc. FFMI 14 workshop at EDM, London (2014)

# Video Performances Juxtaposing STEM with Creativity

Elizabeth J. Hartnett[1], Nils Malzahn[2], and Rick Goldsmith[3]

[1]Open University, Milton Keynes, UK
eliz.hartnett@open.ac.uk
[2]RIAS e.V, Duisburg, Germany
nm@rias-institute.eu
[3]Catchermedia, Herefored, UK
rick@catchermedia.co.uk

**Abstract.** Creative performance through participative video making is a means to engage students in STEM subjects by arousing their curiosity. This paper describes how to engage students in a video making process to support enhanced reflection and understanding of specific topics of study.

**Keywords:** performance, participative video, juxtapose, STEM, creativity.

## 1 Supporting STEM Learning by Creative Performance

Young people are insufficiently motivated to take STEM (Science, Technology, Engineering & Maths) subjects at school or university. Intrinsic motivation is created through a desire to know and to have an interest in something. The EU project JuxtaLearn aims to foster learning by stimulating curiosity and supporting learners through performative and productive activities. Students' performance is substantiated in the form of creative, participatory video creation, together with sharing and commenting of videos in a learning community. Participatory video (PV) involves an inquiry process in which participants are asked to juxtapose standard interpretations with playing with concepts, exploring them creatively [1] and learning in groups. PV provides a medium in which to provoke conflict that facilitates learning [2] by drawing students' attention to discrepancies and opposing viewpoints between their different interpretations of a topic. This encourages iterations between performance and editing activities as students clarify shared understanding, and as the teacher advises them. Hence, PV provokes and encourages reflection [3]. The whole learning pathway is orchestrated into eight stages: (1) identify a subject, (2) demonstrate a standard teaching activity (STA) by the teacher, (3) interpret the STA in groups, (4) perform and create a video, (5) compose the final video, (6) share the video, (7) discuss the videos, (8) review. As creativity is not a widespread approach to teaching, there is a need for a performance framework that helps students to interpret and understand their achievements, encouraging them to compare their performance with other performance on similar topics, and to reflect on how their audience perceives their work. While students are planning for and working on their specific video performance, the teachers are provided with awareness tools to enable them to help all groups,

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 570–571, 2014.
© Springer International Publishing Switzerland 2014

i.e. coordinating students where necessary, helping with topic related questions and keeping them motivated and engaged. To support the student's learning we provide two palettes: a creative performance palette and a practical performance palette. Students must use the creative performance palette (e.g.: a pack of prompt cards) to start crafting their performance, discussing and planning. The practical performance palette supports the students in juxtaposing their creative performance with their teacher's STA, e.g. a classroom experiment, by providing reminders, lists and checkpoints. The process of producing a performance has five steps (within stages 4-6): development (idea generation; initial planning of scenes; dramatic script), pre-production, production, post-production and distribution. A creative performance palette encourages the students' creativity at the development and post-production steps, whereas a practical performance palette structures the creation process through check lists, with constant reference back to a storyboard. The creative process runs throughout the discussion of their video performance enabling differing perspectives on similar STEM topics and offering opportunities for understanding. Besides the performance palettes, the storyboard acts as a road map to help students create a shared language and providing a solid foundation on which to place the components of the video. Not all students will have the same understanding of the topic or concept under study; understanding comes through discussion that these objects instigate, and the opportunity that video making gives to juxtapose performance with the STA. We support students in creatively exploring a STA through metaphors and juxtaposed understandings, providing students with opportunities to compare and contrast, producing structural comparison that initiates a deeper learning [4]. When students compare their performance with the STA, they have opportunities to note and reflect on subtle differences, a process that might lead to serendipitous insights and lateral thinking. Using imagination to define problems facilitates transformative learning [1], and verbalizing the learning through participating and discussing is intrinsically rewarding [5]. First trials at a UK school and Spanish university show that the JuxtaLearn approach not only stimulates targeted teaching for the students, but it also helps to identify the source of students' misunderstandings. However, the students had to be reminded to focus on the subject matter instead of spending too much effort on the performance (e.g. effects or settings).

## References

1. Mezirow, J.: Transformative learning: Theory to practice. New Directions for Adult & Continuing Education 74, 5 (1997)
2. Jones, A.: Constructivist learning theories and IT. Sage in association with the Open University, City (1995)
3. Forsyth, R., Carroll, K., Reitano, P.: Illuminating everyday realities: the significance of video methods for social science and health research. International Journal of Multiple Research Approaches 3(3), 214–217 (2009)
4. Christie, S., Gentner, D.: Where hypotheses come from: Learning new relations by structural alignment. Journal of Cognition and Development 11(3), 356–373 (2010)
5. Bandura, A., Grusec, J.E., Menlove, F.L.: Observational Learning as a Function of Symbolization and Incentive Set. Child Development 3, 499 (1966)

# Eclipse Student (In)Activity Detection Tool

Sokratis Karkalas and Sergio Gutierrez-Santos

Department of Computer Science and Information Systems,
Birkbeck, University of London,
Malet Street, London WC1E 7HX, United Kingdom
{sokratis,sergut}@dcs.bbk.ac.uk
http://www.dcs.bbk.ac.uk/

**Abstract.** This paper presents a tool that has been developed to support teaching of computer programming in practical laboratory sessions. The tool has been implemented as a plugin for the Eclipse IDE and can be deployed in a distributed manner to provide real-time monitoring of the students' activity.

**Keywords:** computer-supported teaching/learning, learning design, learner motivation, distributed and interoperable tools.

## 1 Introduction

The seminal ideas behind this system originated in the MiGen ESRC/EPSRC project [2]. This work is an attempt to extend the use of these tools to the domain of computer programming. A similar project that uses Eclipse to monitor user actions is [3]. This system helps tutors create interactive tutorials. Apart from the purpose there are two major differences between [3] and this tool. The former lacks interoperability support and does not exploit a distributed architecture.

Learning programming in the lab is often hindered by the student's inability to complete the work without help. There are students that due to personal, social or cultural reasons find it difficult to interact with peers and tutors. If students that need help do not ask for it, then it is unlikely to receive it in a timely fashion. Tutors do not know which students need their attention at any given time. On the other hand if students that do not need help do ask for it, then it is very likely to receive much more than what they actually need. These students will not apply enough personal effort to achieve the required task and as a concequence of that the learning outcome for them may eventually be much less than expected.

## 2 The Tool

The tool monitors student activity in the Eclipse Integrated Development Environment (IDE) and consists of the following two main components: *Observer/Watcher* and *User Interface*. The installation process has to be initiated by the

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 572–573, 2014.
© Springer International Publishing Switzerland 2014

student. After the plugin is installed it behaves as an integral part of the Eclipse platform.

*Observer*: A student is considered to be active when she is typing text in Eclipse. The system detects keystrokes and records events in the form of indicators. If the system does not receive keystrokes for more than 5 seconds it records an inactivity indicator. If it is in inactive state and receives a keystroke it records an activity indicator of the same format. From that point on the user is considered to be in active state.

*User Interface*: The purpose of the tool is to help laboratory assistants and lecturers to provide a more focused and efficient service to the students. The tool in that respect has to monitor student activity and communicate that information to the tutors in real time. Similar systems typically provide a client application that is installed on the Lecturer's computer and is used to centrally monitor and/or administer the process. It is thought that this arrangement may not be adequate for University computer laboratories since in this case the tutor must always be moving from student to student in order to provide help. Therefore, the tool was designed to be used in a de-centralised manner through mobile devices. The tool can prioritise the help requests and provide online real-time monitoring of the students' status. It can also generate session statistics and inform tutors about the frequency of help requests per student per machine.

## 3 Discussion and Conclusions

This paper has described the architecture of a system that provides support to tutors in laboratory programming sessions. The system has been tested for stability, fault tolerance and usability in four programming courses during the Spring term of 2014, at the Department of Computer Science and Information Systems, Birkbeck, University of London.

We envision this tool to become an integral part of other computer supported learning environments that have been developed in the department like FLIP [1]. We also plan the next version of the tool to provide real-time monitoring of student progress.

## References

1. Karkalas, S., Gutierrez-Santos, S.: Enhanced javascript learning using code quality tools and a rule-based system in the ip exploratory learning environ-ment (Forthcoming 2014)
2. Mavrikis, M., Geraniou, E., Noss, R., Hoyles, C.: Revisiting pedagogic strategies for supporting students learning in mathematical microworlds. In: Proceedings of the International Workshop on Intelligent Support for Exploratory Environments at EC-TEL, vol. 8, pp. 41–50 (2008)
3. Zhang, Y., Huang, G., Zhang, N., Mei, H.: SmartTutor: creating IDE-based interactive tutorials via editable replay. In: Proceedings of the 31st Inter-national Conference on Software Engineering, pp. 559–562. IEEE Computer Society (2009)

# Map of Relational Links – Visualizing Knowledge Construction

Margarida Lucas

Research Centre Didactics and Technology in Education of Trainers
3810-193 Campus de Santiago, University of Aveiro, Portugal
mlucas@ua.pt

**Abstract.** This paper presents a learning analytic tool for visualizing interactions resulting from discussion posts and replies. The tool is a plugin that adds visualization capabilities to a software that supports the analysis of qualitative data. Its aim is to provide teachers and students with monitoring and feedback of knowledge construction dialogues over time, thus enabling the optimization and personalization of learning processes. Data can be visualized through maps and relational links, allowing users to explore them interactively and chronologically. Analysis parameters were defined by the researcher and different teachers during the design phase of the tool and tests are still being conducted to improve and fine tune the tool's usability.

**Keywords:** learning analytic tool, visualization, knowledge construction, asynchronous discussions.

## 1 Introduction

Asynchronous online discussions play a key role in technology enhanced learning scenarios and are frequently used to promote knowledge construction and higher-order thinking skills. Online discussions, usually in the form of written messages, can reveal students' behaviour during these processes which, from a learning analytics point of view, may help us understand and optimize learning and the environments in which it occurs [1]. To understand learning processes though, we need to consider different aspects that come into play within online learning environments, such as the actions learners take, the representations they use, their engagement in discussions and how such discussions unfold over time. This perspective requires looking at mediating tools, social activity and time as intertwined dimensions of learning [2].

## 2 MRL – Map of Relational Links

MRL is a plugin tool that adds visualization capabilities to the web based qualitative content analysis software webQDA [3], i.e., it builds on data previously coded using this software. MRL offers a basic set of queries which are presented in an interactive visual map format, displaying individual and group interactions over time. Queries

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 574–575, 2014.
© Springer International Publishing Switzerland 2014

include: (i) the list of all posts and comments; (ii) the relation between them (based on sequence of thoughts, the continuity given to the topics covered and the identification of traces that were left throughout the discussions by the participants); (iii) the day and time of the post/comment and (iv) the 5 phases of knowledge construction proposed by the Interaction Analysis Model (IAM) [4]. For each query the user can select a participant and set a time interval, but queries can also be selected as a whole. In Fig. 1 an example for all queries (although only one post regarding query (i) is presented) is shown in a visual map format. Data corresponds to the interactions of 46 participants during 22 days of discussion. Upon closer and detailed scrutiny (possible with the selection of individual queries and the application of different filters) it is possible to envisage aspects relevant to the process of knowledge construction that the IAM or discourse analysis alone cannot provide.

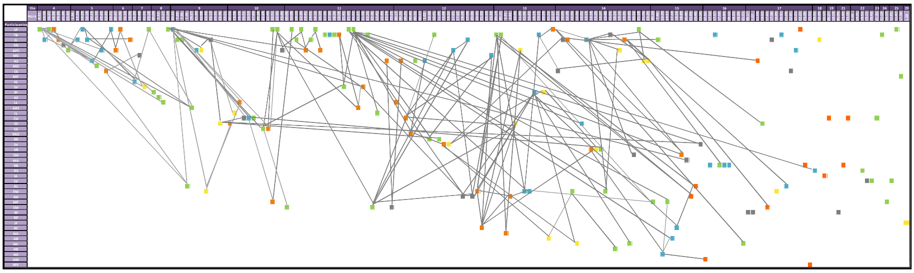


**Fig. 1.** Visual map of online interactions

These relate, for instances, to the emergence and recognition of patterns that may help us identify: (a) actions of social interaction, conflict, negotiation, testing and knowledge construction; (b) the participant(s) who prompted those actions, replied to them or abandoned them; (c) the time actions occurred, how long they endured or evolved to other actions. Hovering over any colour will display the coded message and clicking any relational link will highlight its path, where it began, expanded and ended. Data can be visualized as an integrated whole, but also at an individual level, enabling us to understand the participant's cognitive schema during learning processes and his/her performance and behaviour in relation to the group.

## References


1. SoLAR, Open Learning Analytics: An Integrated &Modularized Platform. White Paper, Society for Learning Analytics Research (2011)
2. Dyke, G., Mayfield, E., Howley, I., Adamson, D., Rosé, C.: Analysis of Discourse and the Importance of Time. In: 1st International Workshop on Discourse-Centric Learning Analytics, LAK 2013 (2013)
3. webQDA, `https://www.webqda.com/?lang=en`
4. Gunawardena, C., Lowe, C., Anderson, T.: Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. J. Edu. Computing Research 17(4), 395–429 (1997)

# The weSPOT Inquiry Space: A Personal and Social Environment for Inquiry-Based Learning

Alexander Mikroyannidis

Knowledge Media Institute, The Open University
Milton Keynes MK7 6AA, United Kingdom
Alexander.Mikroyannidis@open.ac.uk

**Abstract.** This paper presents a personal and social environment for Inquiry-Based Learning (IBL). This environment enables teachers and students to combine a variety of inquiry components and generate mashups that can be adapted in order to meet their personal needs and preferences. This approach also allows users to share their mashups and use them collaboratively with their peers in order to perform their inquiries.

**Keywords:** Inquiry-Based Learning, personalised learning, social learning.

## 1 The weSPOT Inquiry Space

The European project weSPOT[1] adopts a personal and social approach for supporting IBL in secondary and higher education. The aim of this approach is to support students in personalising their IBL environment, as well as in building, sharing and enacting inquiry workflows individually and/or collaboratively with their peers [1]. These principles have driven the development of the weSPOT inquiry space[2], a personal and social IBL environment that reuses and extends the Elgg open-source social networking framework[3].

The development of the weSPOT inquiry space has been primarily based on the following requirements:

- A widget-based architecture enables the personalisation of the inquiry environment, allowing teachers and students to build their inquiries out of mash-ups of inquiry components.
- Students can connect with their peers and form groups in order to build, share and perform inquiries collaboratively.

Inquiries in the weSPOT inquiry space are consistent with the weSPOT pedagogical IBL model [2]. According to this model, an inquiry consists of the following 6 phases: (i) Question/Hypothesis, (ii) Operationalisation, (iii) Data Collection, (iv) Data Analysis, (v) Interpretation/Discussion, and (vi) Communication.

[1] http://wespot.net
[2] http://inquiry.wespot.net
[3] http://elgg.org

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 576–577, 2014.
© Springer International Publishing Switzerland 2014

The weSPOT inquiry space enables its users (teachers and students) to create mashups of their preferred inquiry components, assign them to different phases of an inquiry, share them with other users and use them collaboratively in order to carry out an inquiry. When creating a new inquiry, users are provided with a set of recommended inquiry components for each phase of the inquiry. They can then customise these sets of components by adding, removing and arranging inquiry components for each phase of the inquiry.

The weSPOT inquiry space offers a variety of inquiry components to teachers and students, enabling them to create, edit and share hypotheses, questions, answers, notes, reflections, mind maps, etc. Some of these components communicate with the APIs of REST web services offered by external tools. One of these external tools is the ARLearn mobile app[4], which allows students to collect different types of data (photos, videos, measurements, etc.) with their smartphones and share them with other inquiry members via the weSPOT inquiry space. A Learning Analytics dashboard visualises all the activities taking place within an inquiry, enabling teachers to monitor the progress of their students and students to self-monitor their progress.

Additionally, users have access to external resources and widgets and can use them in their mashups together with the inquiry components offered by the weSPOT inquiry space. These resources and widgets originate from external Learning Management Systems (LMSs), such as Moodle or Blackboard, or from external repositories of widgets, such as the one offered by the European project Go-Lab[5]. In order to integrate external resources originating from LMSs, we have implemented the IMS Learning Tools Interoperability (LTI) specification[6], thus allowing teachers to include in their inquiries either course components from LMSs, such as discussion forums or quizzes, or entire LMS courses.

**Acknowledgments.** The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement N° 318499 - weSPOT project.

## References

1. Mikroyannidis, A., Okada, A., Scott, P., Rusman, E., Specht, M., Stefanov, K., Boytchev, P., Protopsaltis, A., Held, P., Hetzner, S., Kikis-Papadakis, K., Chaimala, F.: weSPOT: A Personal and Social Approach to Inquiry-Based Learning. Journal of Universal Computer Science Special Issue on Cloud Education Environments 19(14), 2093–2111 (2013)
2. Protopsaltis, A., Hetzner, S., Held, P., Seitlinger, P., Bedek, M., Kopeinik, S., Rusman, E., Firssova, O., Specht, M., Haimala, F., Kikis-Papadaki, K., Okada, A., Mikroyannidis, A., Scott, P.: Pedagogical and Diagnostic Framework (2013), `http://portal.ou.nl/documents/7822028/3b2c3110-98bf-4686-b2bd-5756fcab93f1`

[4] `http://portal.ou.nl/en/web/arlearn`
[5] `http://www.golabz.eu/apps`
[6] `http://www.imsglobal.org/lti/index.html`

# A Framework to Adapt Gamification in Learning Environments

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

[1] Université de Lyon, CNRS, INSA-Lyon,
LIRIS, UMR5205, F-69621, France
baptiste.monterrat@insa-lyon.fr
[2] Magellan, IAE Lyon, Université Jean Moulin Lyon 3, France
elise.lavoue@univ-lyon3.fr
[3] LUNAM Université, Université du Maine,
EA 4023, LIUM, 72085 Le Mans, France
sebastien.george@univ-lemans.fr

**Abstract.** Many learning environments are quickly deserted by the learners, even if they are efficient. Gamification appears as a recent game-based learning approach to enhance the learners' motivation. The difficulty with this approach is that people have various expectations from games, and react differently face to specific game mechanics. In order to adapt the game mechanics of the developed game elements, we propose a player model complementary to existing learner models. This model aims to predict to which game mechanics the user is responsive, and is used to adapt the gamified features of the system.

**Keywords:** Gamification, Adaptation, Motivation, Player Model.

## 1 Introduction

Technology-enhanced learning systems are becoming more and more efficient, especially by taking into account the learning goals of the activity and the knowledge model of the learners. However learning activities are not always motivating for the learners and do not prevent them from taking off. Simultaneously, research in the field of game-based learning aims at making the learning activities funnier and more engaging for the users by proposing two main approaches: learning games and gamification. Learning games refer to the use of games for learning purposes; gamification relies only on game design elements [1]. Turning the learning environment into a serious game requires a complete redesign, which could be very expensive and time consuming. In this paper, we are interested in gamification in order to integrate gaming features in already existing learning environments.

People do not have the same emotional responses to game mechanics (e.g. some are highly responsive to competition, whereas others are not). This paper aims to propose a model and a process to raise motivation for learning environments that are not intrinsically motivating, by gamifying them in a strongly adaptive way.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 578–579, 2014.
© Springer International Publishing Switzerland 2014

## 2 Adaptation of Gamification with a Player Model

How to take into account the player types diversity? Adding gaming features for all the existing player types would entail a high risk of overloading the user interface with too many buttons. So far, a common approach to fulfill the player's expectations with gamification is to use the game elements that generally work the most (e.g. points, badges, leaderboards). But learners are usually not engaged with all these elements. Therefore, this diversity among learners-players requires a radical change in game mechanics for a game to be strongly adaptive. Thus we decide to select the game elements according to a player-types model. For instance, the players interested in socializing will be provided with a "share" button to post on social networks, and the players interested in competition will be provided with a leaderboard [2].

The role of the adaptation engine is to find the most relevant gaming feature for a user, according to the player model. The gaming features are represented by a vector based on the classification of Ferro *et al.* [3] with values between 0 and 1. For instance, the leaderboard vector could be [*dominant 1, objectivist 0.4, humanist 0.2, inquisitive 0, creative 0*]. The users are represented in the same way. Indeed, users are generally interested in more than one game mechanic. When gaming features have been designed, the values for their player types vector are set according to the game mechanics they implement. Unlike the game mechanics vector, the player types vector is dynamic: it changes according to the user's interactions with the system.

## 3 Conclusion

Most existing game adaptations focus on adjusting the difficulty level of the game, while gamification requires real changes among game mechanics to support motivation. We argue that, in the context of gamified learning environments, the difficulty has to be handled by the learner model, and the gamification system should focus on adapting the game mechanics. Accordingly, we propose a gamification layer composed of gaming features, a player model, and an adaptation engine for selecting the features for the users.

## References

1. 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, pp. 9–15 (2011)
2. Monterrat, B., Lavoué, É., George, S.: Motivation for learning: Adaptive Gamification for Web-Based Learning Environments. In: Proceedings of the 6th International Conference on Computer Supported Education (CSEDU), pp. 117–125 (2014)
3. Ferro, L.S., Walz, S.P., Greuter, S.: Towards personalised, gamified systems: An investigation into game design, personality and player typologies. In: Proceedings of the 9th Australasian Conference on Interactive Entertainment: Matters of Life and Deathp, p. 7 (2013)

# Design and Operationalization of Learning Situations Based on Patterns for a Public in Professional Development

Lahcen Oubahssi, Claudine Piau-Toffolon, Jean-Pierre Clayer, and Nour El Mawas

Laboratoire d'Informatique de l'Université du Maine (LIUM), EA 4023
LUNAM Université, Université du Maine
72085 Le Mans, France
Surname.Name@lium.univ-lemans.fr

**Abstract.** The research presented in this paper aims to support teachers design learning scenarios with a context-sensitive tool exploiting a library of patterns. This research was carried out in partnership with PARTAGE an association that offers professional training for jobseekers in a design-based research approach. The goal is to assist the association's pedagogical team in the expressing and the formalization of their learning situation through pedagogical scenarios in order to promote the capitalization and the reuse among their community. Moreover, we defined an engineering design process framework and we refined it in collaboration with trainers. To conduct this study, we analyzed the process guiding the jobseeker's courses and we formalized the training practices as design patterns. We also studied the operationalizing feasibility of these patterns in Learning Management System.

**Keywords:** Technology-Enhanced-Learning Systems, Pattern, Pedagogical Scenario, Operationalization, Professional development, Training Development, Design Activity, Instructional Design.

## 1 Introduction

Many models, methods and design tools have been proposed to analyze, implement, and operate models for learning scenarios. Learning Scenarios are designed according to two approaches types. The first one is a well-known modeling approach based on EML (Educational Modelling Language), as the IMS Learning Design specification leads to enable the design of computational models (in the meaning of understandable by a computer) which could be enacted by compliant systems. However these specifications are not really usable by teachers and do not enforce design processes that support the creation of pedagogically sound designs. The second approach allows designers (may be assisted by modeling specialists) to define their own EML, by specifying a domain-specific language and by using it for building their scenarios while ensuring the possibility of translating this specific language to a standard and operationalizable one. We chose to base our research on this second approach because it

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 580–581, 2014.
© Springer International Publishing Switzerland 2014

exceeds the expressiveness limits of the previous one. We propose a teacher-centered design approach in which teachers are able to build their learning scenarios by the use of patterns. We intend to help teachers to formalize their educational needs in the form of patterns that can then be operationalized in Technology-Enhanced-Learning (TEL) systems without any assistance from pedagogical engineers or experts.

## 2 Our Contribution

According to the design-based research methodology we defined an engineering design process framework and refine it with association trainers. In a first time, we analyzed the pedagogical practices of the PARTAGE association. In a second step, we identified the different phases of the support process of a jobseeker. Finally, we carried out the pedagogical practices of the association on the basis of a questionnaire. The analysis of this survey allowed us to identify and describe the different phases of the jobseekers's course. This analysis of pedagogical practices of PARTAGE helped us to define the objectives of the association, identify the problematic, and propose solutions (patterns) to improve existing scenarios.

We note that the main objective of this contribution is to help the trainers to control the mediated learning situations. For this purpose, we propose tools and methods to assist them all along their instructional design process. This research builds on the work of the educational re-engineering of Technology-Enhanced-Learning systems, a pattern-based visual approach, and learning scenarios operationalization. Our aim is to help association trainers to express, formalize, and operationalize their learning scenarios through patterns. Trainers of the association PARTAGE realized patterns of a scenario with the editing tool. Then we generate the pattern in XML format via the patterns editor and then map concepts of the scenario with those of the platform. Main outcomes of the conducted study with PARTAGE consists in (1) a design approach based on patterns, (2) an editing tool which allows trainers to express their learning scenarios according to their own pedagogical way of designing, and (3) an operationalization process.

## 3 Conclusion

In this paper, we address the problem of instructional design in the field of the professional training. We present a solution allowing trainers to design and operationalize learning situations from patterns according to their own design approach identified as pedagogical design scheme. At the end of a pilot study, we asked two trainers to evaluate the tool and the patterns produced. Despite the lack of user friendliness of the first version prototype, they have appreciated the opportunities of formalization of the pedagogical scenario and the expressiveness of the artifacts resulting from the editing session: « *this approach forces us to be more rigorous* ».

**Acknowledgement.** This work and submission are funded by the French GraphiT project [ANR-2011-SI02-011] (http://www-lium.univ-lemans.fr/~laforcad/graphit/).

# Digital Technologies, Teachers' Competences, Students' Engagement and Future Classroom: ITEC Project

Neuza Pedro, João Filipe Matos, and Ana Pedro

Institute of Education- University of Lisbon, Portugal
{nspedro,jfmatos,aipedro}@ie.ul.pt

## 1 iTEC Project

iTEC -Innovative Technologies for Engaging Classrooms- is a four-year project in which European Schoolnet is working with education ministries, technology providers and research institutions to transform the way that technology is used in teaching and learning. With 26 project partners, including 14 Ministries of Education, iTEC is the largest initiative yet on the design of learning and teaching for the future classroom. Within iTEC, educational tools and resources have been piloted in over 2,000 classrooms across 19 European countries with the key objective of providing a sustainable model for fundamentally redesigning teaching and learning.

Started in 2010, education researchers, teachers, education ministries, IT providers representatives and other specialists in the field of education worked together to design the future classroom. The project aims to develop more meaningful visions and scenarios for the future classroom by putting in place a user-centered design process and rigorous testing methodology. All learning activities and new webtools designed for the future classroom are co-developed with teachers and are validated in large-scale pilots (five overlapping 18-month piloting cycles) in order to determine whether they can have the potential to be widely adopted by schools in Europe. The project as developed and piloted three innovative educational tools:

- Eduvista, a toolkit for designing and sharing Future Classroom Scenarios (http://eduvista.eun.org/);
- Edukata, a toolkit for Innovative Learning Activity Design (http://edukata.fi/). It was created upon a collaborative design process that allows educators to design future classroom activities;
- Eduteka: the technologies for Advanced Learning Activity Design (http://itec.eun.org/web/guest/eduteka). It has been developed and tested with the involvement of teachers and students involved in the classroom pilots. The tool also demonstrates 'intelligent' advice on resources, with the capability to make personalized, informed recommendations, based on the teacher's local context.

## 2 Teacher Skills and Competences for Classrooms of the Future

iTEC does not just focus on diffuse futuristic visions; instead, it provides educators with the necessary learning resources and pedagogical tools to allow them to innovate within their teaching and learning practices. However, in order to design the future

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 582–583, 2014.
© Springer International Publishing Switzerland 2014

classroom, iTEC Project saw as required the need to identify the technical and pedagogical skills that teachers will need to effectively act in the classroom of tomorrow. Therefore, iTEC learning activities and webtools were designed considering an ICT Competence Framework for Teachers [2]. This framework is aligned with (i) today's key technological and educational trends and with (ii) the skills that are expected to be evidenced by learners in the near future, which are frequently referred to as 21st-century skills. This framework considers six different domains of teachers work: 1. Understanding ICT in education, 2. Curriculum and assessment, 3. Pedagogy, 4. ICT (digital literacy skills), 5. Organisation and administration, and 6. Teacher professional learning.

## 3 Evaluation of the Impact of iTEC Pilots in Schools

From 2010 to 2014, the project conducted five cycles of design and testing of learning activities. The evaluation of iTEC pilot activities draws information from several sources including questionnaires and case study data collection which includes lesson observations and interviews with the teachers, head teachers and students.

The report of the results founded so far particularly focus on students' learning outcomes and engagement, teachers' digital competence and pedagogical use of ICT. In it, four out of five teachers stated that students had become more deeply engaged in their work, and that the iTEC pilots had positively impacted on students' attitudes to learning. Over 80% of teachers also agreed that the pilot enabled the students to engage in active and independent learning, and that the Learning Activities provided students with new ways to express ideas. Teachers also felt that the pilot had a positive impact on their own attitudes and practices relating to ICT; 79% of the 826 teachers surveyed replied that their knowledge of ICT was improved through taking part in the pilot, with 84% indicating that they intended to use ICT more frequently in future. Above all, 87% of participating teachers felt that using iTEC Learning Activities enabled them to incorporate new pedagogical practices into their classroom activities [1].

The project results are likely to be transferred throughout the participating school and it presents the potential to be taken to scale in order to achieve European ICT-driven economic growth policies as the ones identified in the Horizon 2020 programme.

## References

1. Lewin, C.: Evaluating the impact of iTEC learning activities in schools. iTEC Magazine (1) (2013), `http://fcl.eun.org/c/document_library` (retrieved)
2. UNESCO: ICT Competency Framework for Teachers (2011), `http://unesdoc.unesco.org/images/0021/002134/213475e.pdf` (retrieved)

# Heterogeneous Educational Data Integration and Knowledge Discovery to Supporting Competency Assessment in SCALA Web Tool

Alex Rayón, Mariluz Guenaga, and Asier Nuñez

DeustoTech Learning – Deusto Institute of Technology – University of Deusto
Avda. Universidades 24, 48007 Bilbao, Spain
{alex.rayon,mlguenaga,asier.nunez}@deusto.es

**Abstract.** The lack of data interoperability among different educational systems imposes a challenge to data analytics. To face these problems, we have developed SCALA (*Scalable Competency Assessment web platform through a Learning Analytics approach*), an integrated analytics system that employs Learning Analytics techniques to visualize in a single interface enriched indicators to teachers and learners, gaining insights into their habits and the impact of their learning activities.

**Keywords:** learning analytics, competency-assessment, learning metrics, dashboard.

## 1 Introduction and Related Work

ICT are becoming increasingly important learning mediating tools. Accordingly, there is currently a growing interest in better exploiting data from various sources to help universities to be more effective. This interest has led to the emergence of the area of research and application called Learning Analytics. However, the lack of data interoperability among different data systems imposes a challenge to data mining and analytics that rely on diverse and distributed data. To face these problems, we have designed and developed SCALA (*Scalable Competency Assessment web platform through a Learning Analytics approach*), an integrated analytics web system that processes datasets from a wide variety of educational web technologies.

Following, and based on Chatti [1] we describe a reference model for LA based on four dimensions and identify various challenges and research opportunities in the area of Educational Data Mining in relation to each dimension. *1) What?* According to IMS [2], in the current state of online curriculum delivery, there are a non-standard and/or completely non-existent set of metrics for measuring learning activity content, sourced from an increasingly wider spectrum of solutions and providers; *2) Who?* Few studies involve the teacher by supporting the monitoring of students' activities [3] or the learner by providing feedback [4]; *3) Why?* Assessment and feedback are less common objectives in the reviewed literature. Several methods have been applied for assessment and feedback but the most common are statistics and visualizations; 4) *How?* Based on Baker and Siemens [5], we can establish three categories: Prediction methods, Structure discovery, and Relationship mining.

C. Rensing et al. (eds.): EC-TEL 2014, LNCS 8719, pp. 584–585, 2014.
© Springer International Publishing Switzerland 2014

## 2 SCALA and a Case Study at the University of Deusto

SCALA is a gathering and analysis system that can be described in three stages: observation, analysis and intervention [6]. During the observation phase, actions performed by the learner are collected and stored; this data is filtered and processed. During the analysis phase, the data is analyzed to detect patterns that help define the current situation of teachers and learners. Finally, the data is showed in a dashboard to help them reflect on their performance given the current situation. SCALA, is focused on the three stages. In the observation one, with the ETL process that through a set of web services gathers data from different educational web technologies. In the analysis step, we use three different statistical algorithms to detect and extract knowledge from the large datasets. And, lastly, in the intervention step, we provide awareness by displaying visualizations to the participants of the learning activity. The specific objectives that contribute to reach this aim are: a systematic process of educational data analysis for online courses (what we will call SCALA-process); the definition of a suitable model of educational trace analysis (the so-called SCALA-model); and the design and implementation of a monitoring and visualizing web platform (the extensible web system called SCALA-platform).

Over 2013/14 academic course, a series of experiments have taken place at the University of Deusto. 200 students of five different subjects (Macroeconomics, Project Management Office, Programming, Physics and Entrepreneurship) have participated. The experiment was carried out in two phases: at first, students were informed about the enriched assessment rubrics during a two hour lecture session. The Specific emphasis was given to the interaction analysis indicators which had to be incorporated into the classic assessment rubrics. In the second phase, the students resolved the different activities over the educational web technologies that will later be described. The first version of SCALA extracts evidences from a set of educational web technologies that have different data formatting (Moodle, Google Docs, Twitter, MediaWiki and Youtube). From all these heterogeneous data sources, we take the author and the timestamp, to be able to perform further analysis.

In the future, we would like to be able to assess more competencies such as Entrepreneurship, Creativity or Problem-solving, if we consider more educational technologies such as Serious Games, eBooks or OER repositories.

## References

1. Chatti, M.A., Dyckhoff, A.L., Schroeder, U., Thus, H.: A reference model for learning analytics. International Journal of Technology Enhanced Learning, IJTEL (2011)
2. IMS. Learning measurement for analytics whitepaper. IMS Global Learning Consortium (2013)
3. Bakharia, A., Dawson, S.: 2010. SNAPP: A bird's eye view of temporal participant interaction. Learning Analytics & Knowledge (2010)
4. Clow, D., Makriyanni, E.: Ispot analysed: Participatory learning and reputation. In: Proceedings of the 1st International Conference Learning Analytics & Knowledge (2010)
5. Baker, R., Siemens, G.: Educational data mining and learning analytics (2013)
6. Elias, T.: Learning analytics: Definitions, processes and potential (2011)

# The Social Navigator: A Personalized Learning Platform for Social Media Education

Michael Schmidt[1], Uta Schwertel[1],
Christina Di Valentin[2], Andreas Emrich[2], and Dirk Werth[2]

[1] Information Multimedia Communication AG, 66123 Saarbrücken, Germany
{michael.schmidt,uta.schwertel}@im-c.de
[2] DFKI, Institut für Wirtschaftsinformatik, 66123 Saarbrücken, Germany
{christina.divalentin,andreas.emrich,dirk.werth}@iwi.dfki.de

**Abstract.** We present the Social Navigator, an open, Web-based learning platform supporting training of Social Media skills in vocational education. It provides access to a rich pool of resources that enable teachers to integrate Social Media education into the teaching process and offers information and training material to trainees. Designed in an interdisciplinary approach, the Social Navigator builds upon a Social Web skill model, which is used to provide personalized, guided access to resources. Complementary, the integration of Social Web channels stimulates social interaction and helps users to implicitly acquire Social Web skills.

**Keywords:** Social Media Education, Personalized Recommendations.

## 1 The Social Navigator

Social media have become an integral part of everyday's working life, particularly in the ICT industry. The skills needed to professionally use channels such as wikis, blogs, and networking sites include search, selection, comparison, administration, and publication of information. With the goal to increase the awareness for opportunities and risks of Social Media and to convey required skills, we are developing the *Social Navigator* as a personalized and social learning system [1].

For *teachers and trainers*, the system shall motivate the importance of Social Media skills for their trainees' working life and provide support in designing and running projects that convey Social Media skills; for *trainees*, it constitutes a platform that helps acquire Social Media skills through self-assessment tests, personalized learning object recommendations, and dedicated Web-based trainings. Designed as a Social Network, it stimulates communicative exchange such as sharing of experiences made with Social Media in vocational education.

**Social Media Skill Model.** The Social Navigator is centered around learning objects and information about Social Media skills, which describe the capabilities of adequately applying social media technologies [2,3]. In [4] we presented an integrated model describing knowledge and skills for the use of the Social Web for learning and knowledge exchange. At its core, the model identifies four complementary facets for professional Social Media use, namely skills required

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 586–587, 2014.
© Springer International Publishing Switzerland 2014

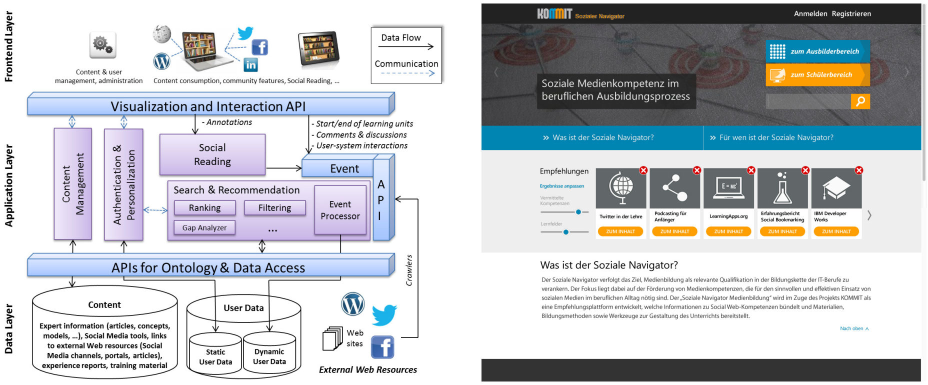



**Fig. 1.** (l) Social Navigator architecture; (r) Screenshot incl. recommendation stream

to *create and edit information*, *communicate and comment on information*, *comprehend and evaluate information*, and *select and manage information*. The skill model is tightly integrated into the Social Navigator, serving as a means for classification, navigation, and recommendation of learning objects.

**Architecture, Design, and Implementation.** The architecture of the Social Navigator follows a three-tier approach (cf. Figure 1(l)), in which learning object metadata as well as dynamic user behavior are used by an ontology-based search and recommendation engine [1], to provide personalized information to end users (see e.g. the recommendation stream skteched in Figure 1(r)). In order to support mobile learning, the visualization and interaction API follows a responsive design paradigm that delivers output to desktop, laptop or tablet PCs in an optimized way. Another central component is a module for Social Reading, which allows users to annotate text fragments in learning objects, add comments or links to these annotations, and share them with others, thus enabling community-driven communication around integrated learning objects.

**Acknowledgments.** This research has been co-funded by the European Social Fund and the German Federal Ministry of Education and Research within the project *KOMMIT*.

## References

1. Schmidt, M., Di Valentin, C., Emrich, A., Schwertel, U., Oloff, C., Kammerer, Y.: A Social and Personalized Learning Platform for Vocational Social Media Education. In: Multikonferenz Wirschaftsinformatik, pp. 820–832 (2014)
2. Baacke, D.: Medienkompetenz–Begrifflichkeit und sozialer Wandel. In: Medienkompetenz als Schlüsselbegriff, pp. 112–124 (1996)
3. Parola, A., Ranieri, M.: Media Education in Action: A Research Study in Six European Countries, vol. 110. Firenze University Press (2010)
4. Oloff, C., Kammerer, Y., Gerjets, P.: Theoretisches Rahmenmodell zu Fertigkeiten bei der Nutzung des Social Web für Wissenserwerb, Wissenskonstruktion und Wissensaustausch. In: 14. Fachtagung Pädagogische Psychologie der Deutschen Gesellschaft für Psychologie, Hildesheim (2014)

# Assessing and Training Social Media Skills in Vocational Education Supported by TEL Instruments

Uta Schwertel[1], Yvonne Kammerer[2], Clara Oloff[2], Peter Gerjets[2], and Michael Schmidt[1]

[1] Information Multimedia Communication AG, 66123 Saarbrücken, Germany
{uta.schwertel,michael.schmidt}@im-c.de
[2] Knowledge Media Research Center, 72076 Tübingen, Germany
{y.kammerer,c.oloff@iwm-kmrc.de,p.gerjets}@iwm-kmrc.de

**Abstract.** The use of social media has become popular in the professional IT sector, requiring specific skills to exploit their full potential. Based on our model on relevant social media skills we report on a study testing related proficiencies of IT vocational students. From this we derive three technology-enhanced training instruments to foster the skills in vocational education processes.[1]

**Keywords:** social media education, skill model, skill testing, TEL instruments.

## 1 Assessing Social Media Skills

Social media, such as blogs, online forums, wikis, and social networking sites, increasingly become an integral part of young adults' personal as well as school or working life. Particularly in the IT sector the use of social media has become popular for both knowledge consumption and production. A competent and reflective use of social media, however, not only requires technical skills (e.g., how to set up a wiki page) but even more a set of cognitive and social skills for learning and knowledge exchange. On the basis of existing models about general ICT literacy (e.g., [1]) as well as models from educational psychology about information problem solving on the internet (e.g., [2]) and collaborative problem solving [3] we derived four skill facets that form the basis for an adequate receptive or productive use of social media in educational or business contexts: *A. select and manage information*, *B. comprehend and evaluate information*, *C. communicate and comment on information*, *D. create and edit information*. The facets are part of our more comprehensive skill model [4].

In order to test students' proficiency regarding the identified skill facets we developed a set of 10 exemplary tasks. For example in one task (testing facet B) students were asked to read a blog article about health risks of laser printers and to argue whether or not they would use this website for a school assignment. All tasks required

[1] The research described in this paper has been co-funded by the European Social Fund and the German Federal Ministry of Education and Research within the project KOMMIT.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 588–589, 2014.
© Springer International Publishing Switzerland 2014

free text entry and were presented on individual laptops in a classroom. Table 1 summarizes the results of our study, cf. [4]. Dark grey areas mark the percentage of students who achieved at least half of the maximally possible points according to our detailed assessment matrix. The results indicate training needs for facets A, B, and D.

**Table 1.** Study results (Sample: 124 vocational students from IT sector, M=20.4y, 93.5% male)

| | | Points (P) achieved wrt. assessment matrix | | | | | | | | | | max. P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Task (T)** | **Skill** | **0** | **0.5** | **1** | **1.5** | **2** | **2.5** | **3** | **4** | **5** | **6** | |
| 1. Point out strategies to keep oneself up to date | A | 25% | | 51% | | | 24% | | | - | - | 4 P |
| 2. Evaluate credibility of blog article | B | 21% | 39% | | 30% | | 10% | | | | - | 5 P |
| 3. Identify inappropriate Facebook posts | C | 8% | | 8% | | 14% | | | 70% | | | 6 P |
| 4. Anticipate problems when collaboratively writing a Wiki article | D | 15% | | 50% | | 31% | | | 4% | | | 6 P |

## 2 Training Social Media Skills

Based on the skill model and the results of our study we are currently realizing three different technology-enhanced training instruments to foster the identified skills.

**Web-Based Training Modules for Students.** We are developing two web-based self-study training modules (20-30 min learning time) which train the skills "Keep oneself up-to-date" (skill of facet A) and "Evaluate the quality and credibility of information" (skill of facet B). For each module we have defined a set of learning goals (e.g., separate relevant from irrelevant information, evaluate author's expertise, etc.) and train them using real world IT relevant scenarios (e.g., "Bring your own device").

**Social Navigator.** We are developing the Social Navigator, cf. [5], an open online resource and collaboration platform, which provides internal and external free educational resources about the topic social media education for teachers, trainers, as well as students of vocational IT education.

**Open Online Course for Teachers.** We are currently devising an open online course for teachers and trainers of vocational IT education, which guides them how to integrate and train the respective skill facets into the classroom or workplace.

## References

1. Katz, I.R.: Testing information literacy in digital environments: ETS's iSkills assessment. Information Technology and Libraries 26, 3–12 (2007)
2. Brand-Gruwel, S., Wopereis, I., Walraven, A.: A descriptive model of information problem solving while using internet. Computers & Education 53, 1207–1217 (2009)
3. Hesse, F.W., Buder, J., Sassenberg, K.: A framework for teachable collaborative problem solving skills. In: European Conference on Educational Research (ECER), Berlin (2011)
4. Kammerer, Y., Oloff, C., Gerjets, P.: Erfassung von Fertigkeiten im Umgang mit dem Social Web. In: Mayer, A.-K. (ed.) Informationskompetenz im Hochschulkontext – Interdisziplinäre Forschungsperspektiven. Pabst, Lengerich (in press)
5. Schmidt, M., Di Valentin, C., Emrich, A., Schwertel, U., Oloff, C., Kammerer, Y.: A social and personalized learning platform for vocational social media education. In: Kundisch, D., Suhl, L., Beckmann, L. (eds.) Proceedings of MKWI 2014, February 26-28, pp. 820–832. University of Paderborn, Deutschland (2014)

# Negotiation-Driven Learning

Raja M. Suleman, Riichiro Mizoguchi, and Mitsuru Ikeda

School of Knowledge Science,
Japan Advanced Institute of Science & Technology, Nomi, Ishikawa, Japan
{suleman,mizo,ikeda}@jaist.ac.jp
http://www.jaist.ac.jp

**Abstract.** Open Learner Models mostly employ positional-based negotiation mechanisms which confine the role of negotiation as a "problem-solving" technique. Negotiation provides an excellent opportunity to learn and understand why a learner holds a certain belief and why these beliefs contradict with that of the system. This study aims at proposing a new paradigm of learning that uses negotiation coupled targeted responses to motivate a learner and enhance their metacognitive skills along with their cognitive skills.

**Keywords:** Negotiation, Metacognition, negotiation-driven learning, interest-based negotiation, learner motivation.

## 1 Introduction

Intelligent Tutoring Systems (ITS) use personalization and user modeling to provide adaptive learning. ITSs employ a Learner Model (LM) to infer the learner's knowledge state which drives this adaptive interaction. Much of the research in ITSs has primarily focused on generation and utilization of LMs. Open Learner Models (OLM) were introduced with the idea to open the LM to the learner which provides them with an opportunity to externalize their knowledge. OLMs aim to increase learner involvement and promote their cognitive & metacognitive abilities through guided content, externalization, scaffolding and negotiation [1]. However, the role of negotiation has been undermined by how it has been used in OLMs. This paper proposes the paradigm of Negotiation-Driven Learning (NDL) with the aspiration to enhance the role of negotiation as a "problem-understanding" technique and use it to promote both cognitive & metacognitive skills of a learner.

## 2 Background

Negotiation helps remove conflicts between the beliefs of a learner and the system which results in a more accurate LM. This not only increases the accuracy of the LM but also promotes metacognitive skills such as reflection and self-monitoring in learners [1]. In recent years a lot of research has focused on the

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 590–591, 2014.
© Springer International Publishing Switzerland 2014

promotion of metacognitive skills in learners. It has been argued that for an Intelligent Tutoring System to be effective, it is necessary that the systems model student's meta-knowledge along with their domain knowledge. Also the fact that not all learners are equally capable of evaluating their knowledge, they should be provided with enough support to externalize and realize their knowledge.

## 3 Negotiation Driven Learning

We aim to take advantage of automated Interest-Based Negotiation (IBN) and behavioral modeling and use them to understand and challenge the motivational state of the learner. Indulge the learner in an active dialog which promotes self-regulated learning, and provide the learner the opportunity to understand "what and how" they learn.

### 3.1 System Architecture

In order to automate the IBN in our system, we propose a model of an Extended Interest-Based Negotiation Agent which extends the computational model described in [2]. We introduce the use of a Behavioral Model (BM) that is used by the system to hold information regarding the behavioral state of the learner. The BM can play a very important role in the IBN process, since for a negotiation strategy to be fruitful, it is imperative that the motivational state of the other party is taken into consideration. Figure 2 shows the architecture of the proposed system.

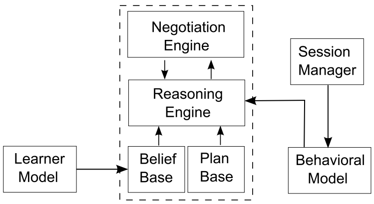



**Fig. 1.** Extended Interest-Based Automated Negotiation Agent

## 4 Conclusion

In this study we have proposed a paradigm of Negotiated-Driven Learning. NDL follows the notion that learning is maximized by participation in the learning process and negotiation provides an excellent opportunity to challenge the learners which promotes metacognitive skills by motivating them to think more objectively about their learning.

## References

1. Bull, S., Pain, H.: Did I Say What I Think I Said, And Do You Agree With Me?: Inspecting and Questioning the Student Model. In: Greer, J. (ed.) AIED 1995, pp. 501–508. AACE, Char-lottesville (1995)
2. Tao, X., Miao, Y., Shen, Z., Miao, C.Y., Yelland, N.: Interest Based Negotiation Automation. In: Huang, D.-S., Li, K., Irwin, G.W. (eds.) ICIC 2006. LNCS (LNBI), vol. 4115, pp. 211–222. Springer, Heidelberg (2006)

# Adaptive Learner Profiling Provides the Optimal Sequence of Posed Basic Mathematical Problems

Behnam Taraghi[1], Anna Saranti[1], Martin Ebner[1],
Arndt Großmann[2], and Vinzent Müller[2]

[1] Graz University of Technology, Münzgrabenstrasse 35/I, 8010 Graz, Austria
{b.taraghi,martin.ebner}@tugraz.at, s0473056@sbox.tugraz.at
[2] UnlockYourBrain GmbH, Französische Str. 24, 10117 Berlin, Germany
{arndt,vinzent.mueller}@unlockyourbrain.com

**Abstract.** Applications that try to enhance learners' knowledge can profit by the creation and analysis of learner profiles. This work deals with the derivation of an optimal sequence of questions by comparing similar learning behaviour of users of a mathematics training application. The adaptation of the learners' clusters to the answers of the revised optimal question sequence improves learning.

**Keywords:** Learning Analytics, Clustering, Adaptive Learner Profiling.

## 1 Classification of Questions

Related works that use explicit [1] and implicit [2] knowledge from the learners, provided motivation for our previous research [3]. In this work, only implicit knowledge on the learning behaviour was used. The dataset was provided by the Android application *UnlockYourBrain*. The addition questions contain about 8 million answers given by 107603 different users and subtraction, multiplication and division around 2 million answers from 50000 different users. The answer to a posed question can be correct (R) of false (W). By this means, eight sorted answer types were defined for questions with up to five answering options: R, WR, W, WWR, WW, WWWR, WWW and WWWW. An eight-dimensional feature space having the occurrence probabilities of each answer type as dimensions, contains one point per question. The K-Means [4] algorithm was applied to classify the addition questions to 13, the subtraction questions to 10, and multiplication and division to 11 clusters. Questions that needed proportionally more attempts to be correctly answered, were clustered together.

## 2 Adaptive User Profiling and Optimal Sequence

A Markov chain was the adequate model for the learning behaviour and application use of one specific user. The learner always gets a question belonging to a questions' cluster (state of Markov chain) and the answer is always one of the eight answer types (also state of Markov chain). The answering act as well as

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 592–593, 2014.
© Springer International Publishing Switzerland 2014

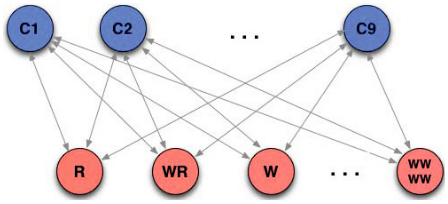



**Fig. 1.** Markov chain model of one user profile

the posing of a new question are both the state transitions within this Markov model. The probabilities of the transitions reflect the answering choices of the user and the question's posing algorithm used by the application (see figure 1). The questions' clusters can undergo a change in number and content due to reclassification. These actions will induce changes in the users' profiles since the latter are coupled per definition with the questions' difficulties clusters.

The ultimate goal of the users' clustering is to identify the learning patterns that are similar (or respectively different) between the learners. Therefore, a hierarchical clustering algorithm [4] was applied to cluster the different Markov chain user profiles; each operation was considered individually. The multidimensional feature space of all user profiles is of size: number of answer types (8) times the number of question's difficulties clusters. A user's Markov chain living in this space, is a point with coordinates that represent the Markov chain's transition probabilities.

When observing the learning behaviour of the users, several sequences of questions could indicate an improvement over time. The sequence plays an important role in enhancing the learning goals of the users if the questions' answers converge over time towards being correct. The application learns which sequence of questions was more beneficial for the user's neighbours from a learning viewpoint. Especially from those neighbours that have matched question-answer pair subsequences, the application can use their further successful progress (if any) to decide and pose the next optimal question.

## References

1. Wei, X., Yan, J.: Learner Profile Design for Personalized E-Learning Systems. In: Proceedings of the International Conference on Computational Intelligence and Software Engineering, pp. 1–4 (2009)
2. Zghal Rebaï, R., Ghorbel, L., Zayani, C.A., Amous, I.: An Adaptive Method for User Profile Learning. In: Catania, B., Guerrini, G., Pokorný, J. (eds.) ADBIS 2013. LNCS, vol. 8133, pp. 126–134. Springer, Heidelberg (2013)
3. Taraghi, B., Ebner, M., Saranti, A., Schön, M.: On Using Markov Chain to Evidence the Learning Structures and Difficulty Levels of One Digit Multiplication. In: Proceedings of the 4th International Conference on Learning Analytics and Knowledge, Indianapolis, USA, pp. 68–72 (2014)
4. Murphy, K.P.: Machine Learning a probabilistic prospective. The MIT Press, Cambridge (2012)

# Teaching and Learning Processes in a School Learning Network

Patricia Thibaut Paez

School of Education, University of Sydney NSW 2006, Australia
patricia.thibautpaez@sydney.edu.au

**Abstract.** Often methods of contemporary schooling assumed teacher-centred practices. I draw on the concept of social learning and learning networks to challenge this model. In particular, I explore the role of a learning technology, a Social Network Site called Edmodo in teaching and learning practices in the primary classroom. Both, a quantitative approach and qualitative technique were utilised. By doing so, I attempt to understand not only the patterns of interaction, but also the underlying meaning of interactions used in the asynchronous environment. This study aims to bridge gaps between contemporary methods of schooling and learning technologies. I will show the applicability of Edmodo in expanding teaching and learning repertoires and in building learning networks in young learners.

**Keywords:** Network learning, primary education, social network sites, social learning.

## 1 Introduction

Current schooling is often critiqued for perpetuating teacher-centered practices. Students and teachers interactions in the classroom frequently play two distinctive roles, with teachers focusing on the delivery of content and students in the appropriation of knowledge. This echoes contemporary schooling expectations of students learning the same content, at the same rate and being assessed under standardized measurements [3]. A major aspect of social learning poses that learning occurs as part of inquiry processes which involves not only enacting the role of students/mentees but also the role of teachers/mentors [2]. As schools are progressively incorporating a blended approach for teaching and learning, it is relevant to see how Information and Communication Technologies (ICT) might influence this educational context. One particular ICT platform where interactions are extended, from face to face mode to an asynchronous mode, are social network sites (SNSs). I examine the use of a SNS with learning purposes called Edmodo in an upper-primary classroom and explore if the inclusion of this learning technology promotes active roles in learners, involving conventional students' interactions and also mentoring processes. In addition, I examine the teacher's role in the online space and the interactions emerging with the inclusion of Edmodo in the learning network of the classroom.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 594–595, 2014.
© Springer International Publishing Switzerland 2014

## 2 Research Context

This study is situated in an Australian independent pre K-12 school located in Sydney that emphasizes innovation in teaching and learning and the use of digital technologies. The study considers the contributions of a year 6 class (11 and 12 years old), composed of 30 students and the head teacher, over four months. Students in this learning environment are equipped with their own laptops/personal devices, a Wi-Fi network and a Learning Management System.

## 3 Method

First, a content analysis technique [5] and the Anderson Schema for tutoring processes was employed [1]. Second, a micro-ethnographic discourse analysis method was employed. This technique allows exploring the meaning underlying the interactions on the Edmodo wall, beyond the quantitative data. Consequently discourse analysis focuses attention on how people adopt tools to communicate and engage in semiotic practices and in turn create social realities.

## 4 Findings

In this analysis, I examine students engaging in teaching practices and the patterns of the teacher's contributions on Edmodo. Both, the content analysis and discourse analysis suggest that students are engaging in rich level of social learning in this environment. Yet, the quantity of interactions by students varied and the nature of their contributions can adopt different forms. Studies on asynchronous learning platforms have evidenced that peer teaching and dialogue can be drivers that increase the quality of students' thinking process [4]. The evidence shows that students are positioning themselves not only in traditional patterns of students' roles but also engaging in mentors' practices, navigating across these two roles. The teacher's contributions are in turn, less significant in directing instruction and rather focused on activities related to the design on instructional activities. Consequently, the data suggest that the use of a social network site with learning purposes leverages the construction of teaching and learning repertoires beyond conventional patterns in the primary classroom.

## References


1. Anderson, T., Rourke, L., Garrison, D.R., Archer, W.: Assessing teaching presence in a computer conference context. Journal of Asynchronous Learning Networks 5(2), 1–17 (2001)
2. Brown, J.S., Collins, A., Duguid, P.: Situated cognition and the culture of learning. Educational Researcher 18(1), 32–42 (1989)
3. Collins, A., Halverson, R.: The second educational revolution: rethinking education in the age of technology. Journal of Computer Assisted Learning 26(1), 18–27 (2010)
4. Johnson, M.G.: Synchronous and Asynchronous Text-Based CMC in Educational Contexts: A Review of Recent. Tech. Trends 50(4) (2004)
5. Krippendorf, K.: Content Analysis: An introduction to its methodology. Sage, Thousand Oaks (2013)

# A Learning Analytics Methodology to Profile Students Behavior and Explore Interactions with a Digital Electronics Simulator

Mehrnoosh Vahdat[1,3], Luca Oneto[1], Alessandro Ghio[1], Giuliano Donzellini[1], Davide Anguita[2], Mathias Funk[3], and Matthias Rauterberg[3]

[1] DITEN - Università degli Studi di Genova, Genoa I-16145, Italy
mehrnoosh.vahdat@edu.unige.it,
{luca.oneto,alessandro.ghio,giuliano.donzellini}@unige.it
[2] DIBRIS - Università degli Studi di Genova, Genoa I-16145, Italy
davide.anguita@unige.it
[3] Department of Industrial Design, Eindhoven University of Technology, 5612AZ Eindhoven, The Netherlands
{m.funk,g.w.m.rauterberg}@tue.nl

**Abstract.** The automatic collection of data concerning the interaction between students and Technology Enhanced Learning (TEL) systems has become increasingly common. Such data availability has led to applications of Learning Analytics (LA) techniques, characterized by the capability of extracting non-trivial patterns from them. In this framework, we describe the methodology we are adopting to explore the way students learn the concepts of digital electronics by exploiting LA. In particular, a simulator named DEEDS (Digital Electronics Education and Design Suite) is used, and we aim at extracting non-trivial knowledge from data, gathered through a logging application, properly realized for monitoring DEEDS usage.[1]

**Keywords:** Learning Analytics, Technology Enhanced Learning, Simulator, Student Interaction.

## 1 Learning Analytics with DEEDS Simulator

Learning Analytics (LA) has developed remarkably in the last decade in the field of higher education by focusing on learning processes [1–3]. By adopting methods of analytics in e-learning, researchers gain more insight into the data collected from students through building models [4]. In the context of TEL, simulation-based environments are efficient tools to improve learning [5]. In this framework, we concentrate on higher education and use DEEDS simulator (Digital Electronics Education and Design Suite) [6] which is used for e-learning in

[1] This work is supported partly by the Erasmus Mundus Joint Doctorate in Interactive and Cognitive Environments, funded by the EACEA Agency of the European Commission.

C. Rensing et al. (Eds.): EC-TEL 2014, LNCS 8719, pp. 596–597, 2014.
© Springer International Publishing Switzerland 2014

digital electronics. We aim to extract precious knowledge via LA methods from data, gathered through a logging application called 'LA Data Collector' (LADC). This study was carried out on the students of Digital Electronics course at the University of Genoa during the spring semester, 2014.

In particular, a comprehensive experiment is designed and around 120 students participated in. An instructional design of the experiment to log the interaction data for comparing to the learning outcomes is performed. This leads us to discover common patterns and cross-correlations between variables for profiling students learning activities and attitudes. The activity data from the whole course by LADC is collected from over 95% of students (see Figure 1). During the experiment, observation, interview, and demographic questionnaire were carried out to acquire the learners' background, motivation, and feedback. A dataset concerning the learners' behavior and their key performance indicators, while interacting with the simulator is being prepared and will be made available as a valuable source for the research community.

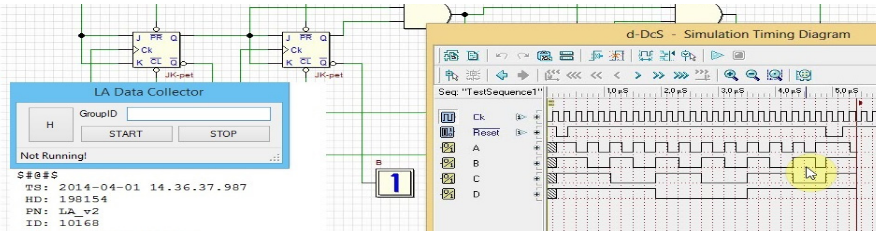



**Fig. 1.** DEEDS simulator used by a student trying to solve an exercise. On the lower left, the log data is shown as well as the LADC interface.

## References

1. Brown, M.: Learning analytics: moving from concept to practice. Educause Learning Initiative (2012)
2. Shum, S.: Learning analytics. UNESCO Institute for Information Technologies in Education. Policy Brief (2012)
3. Bienkowski, M., Feng, M., Means, B.: Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. US Department of Education, Office of Educational Technology, 1–57 (2012)
4. Siemens, G., de Baker, R.S.: Learning analytics and educational data mining: towards communication and collaboration. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 252–254. ACM (2012)
5. Donzellini, G., Ponta, D.: A simulation environment for e-learning in digital design. IEEE Transactions on Industrial Electronics 54(6), 3078–3085 (2007)
6. Ponta, D., Anguita, D., Da Bormida, G., Donzellini, G.: Ten years of activity on computer-aided learning for electronics: Needs, experiences, field evaluation, invited paper by taee 98, iii congreso sobre tecnologías aplicadas a la enseñanza de la electrónica. Universidad Politécnica de Madrid, Madrid, 16–18 (1998)

# Author Index