eduworks

Learning & Student Analytics Conference (LSAC) 2017: Implementation, Institutional Barriers and New Developments

Book of Abstracts

26-27 October , 2017

#LSAC2017

Learning & Student Analytics Conference (LSAC) 2017: Implementation, Institutional Barriers and New Developments - Book of Abstracts

www.lsac2017.org

Edited by:
Dr. Gábor Kismihók, University of Amsterdam
Dr. Ilja Cornelisz, Vrije Universiteit Amsterdam
Dr. Chris van Klaveren, Vrije Universiteit Amsterdam
Alan Berg, University of Amsterdam
Dr. Stefan T. Mol, University of Amsterdam

Publisher: Amsterdam Business School – University of Amsterdam

ISBN/EAN: 978-90-827931-0-9
Financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Programme is gratefully acknowledged.

Contents

Preface

Learning and Student Analytics is slowly and steadily making its way from research to practice. In the past decade, actionable research has been carried out stimulating policy makers and educators to take an ever increasing interest in applying these findings to educational practice. However, despite the available evidence, technology, and many examples of good practices, organisational uptake is slow.

One of the reasons for the slow adoption is the lack of dialogue and cross pollination between core expert groups (policy makers, educational researchers, computer scientists, educational system developers, data miners, infrastructure architects and vendors) in academia and practice. Many stakeholders groups are involved in or affected by Learning Analytics without being aware of it, making the sustainable scaled implementation of learning analytics interventions in practice a challenging endeavour at best. These include educational managers, educational designers, educational policy makers both at the organisational and regional level, student associations, employment agencies, ethics boards, data governance centres, technologists, and so forth. There is a need to involve this wider stakeholder group in this discussion, as they have urgent and substantial claims in this fast growing field.

From a research perspective there is a clear need from the community to maintain and evolve a structured Learning Analytics evidence base. This ongoing exercise also requires out of the box thinking in order to diminish barriers between learning analytics related theories, methods, and the available datasets, in a multidisciplinary environment that is later deployable at scale.

Therefore, the aim of this conference is to bring together researchers and practitioners from a number of disciplines (e.g. education, technology, computer science, management, psychology, economics, IT security etc.), organisational and national policy makers, educational practitioners, students, employers to share and discuss the latest research insights related to learning and student analytics. The conference further provides a platform for stakeholders to engage in critical conversations about current trends and the policy requirements of Learning Analytics.

This conference is organised by the Amsterdam Center for Learning Analytics (ACLA) and the Eduworks-Network.

ACLA (www.acla.amsterdam) is devoted to improving education and labour market outcomes by adopting a comprehensive approach towards learning analytics. ACLA conducts research- and educational activities using and combining insights from information technology and computer sciences, theories of behaviour, learning and education, and rigorous empirical evaluation methods. In doing so, ACLA seeks to provide a fundamental scientific contribution and to structurally improve the quality of education and labour markets for current and future generations.

The objective of the EDUWORKS-Network (www.eduworks-network.eu) is to train talented early-stage researchers in the socioeconomic and psychological dynamics of the labour supply and demand matching processes at aggregated and disaggregated levels. Understanding how the matching process works can prevent mismatches with respect to skills and qualifications, and can lead to an improved balance between the supply of and demand for labour. EDUWORKS brings together researchers from several academic disciplines. namely: Labour Economics, Sociology of Occupations, HRM, Lifelong Learning and Knowledge Management.

Organising and Programme committee:

Alan Berg, University of Amsterdam

Dr. Ilja Cornelisz, Vrije Universiteit Amsterdam

Dr. Chris van Klaveren, Vrije Universiteit Amsterdam

Dr. Gábor Kismihók, University of Amsterdam

Dr. Stefan T. Mol, University of Amsterdam

Prof. Dr. Anwar Osseyran, SURFsara

Contact: LSAC.FGB@vu.nl

www.lsac2017.org

#LSAC2017

We thank our sponsors for their generous contributions:

SURF

Vrije Universiteit Amsterdam

University of Amsterdam

EDUWORKS- Network - EU MSCA Actions

PROGRAMME OUTLINE

October 26, 2017

10.00-10.05 **Word of welcome**

10.05-11.00 **Keynote prof. Dragan Gasevic**

11.00-11.30 **COFFEE BREAK**

11.30-13.00 **Research Presentations**

1. Theory and Methods 1 - Providing student and teacher support
2. Methods and Data 1 - Learning paths and student engagement
3. Data and Theory 1 - Representing digital learning interactions

13.00-14.00 **LUNCH**

14.00-15.00 **Panel discussion - Learning analytics policies**

15.00-15.30 **COFFEE BREAK**

15.30-17.00 **Workshop: LA policy challenges**

1. Workshop 1 - Changing business model of education
2. Workshop 2 - Presenting learning analytics to stakeholders of education
3. Workshop 3 - Learning analytics and organisational culture

19.30-22.30 **CONFERENCE DINNER**

October 27, 2017

10.00-10.05 Word of welcome

10.05-11.00 Keynote prof. Sanna Järvelä

11.00-11.30 COFFEE BREAK

11.30-13.00 Research Presentations

1. Theory and Methods 2 - Matching education, goals and labour market demand
2. Methods and Data 2 - Self-regulated and adaptive learning
3. Data and Theory 2 - Social comparison feedback and motivation

13.00-14.00 LUNCH

14.00-15.00 Panel discussion Cooperation between research and industry

15.00-15.30 COFFEE BREAK

15.30-17.00 Sessions: LA Applications in Education

1. Session 1 - Issues of privacy and ethics
2. Session 2 - Early warning systems and predicting student success or failure
3. Session 3 - Learning Analytics infrastructure and dashboard development

END CONFERENCE

DETAILED PROGRAMME

October 26, 2017

10.00-11.00 Keynote - Dragan Gasevic
Room: A1.03

11.00-11.30 Coffee Break

11.30-13.00 Research Presentations
Theory and Methods 1 - Providing student and teacher support
Room: A1.03, Chair: Alan Berg

Nynke Bos, Leiden University
Using Educational Design Research to Develop Actionable Analytics to Support First Year Students

Jocelyn Manderveld, SURFnet
SURF Learning Analytics Experiment: Hands-on experience for Dutch higher education

Alan Berg , Central Services (ICTS), University of Amsterdam
Should students have a right to analytics?

Methods and Data 1 - Learning paths and student engagement
Room: A2.09, Chair: Ilja Cornelisz

Joel Howell, Curtin University
Mastery vs. Avoidance? Impact of grade, sender, comparative information and message style on student affect and academic resilience

Anouk Gelan, Universiteit Hasselt, University of Amsterdam
Analyzing and visualizing learner behavior with learning analytics in language and mathematics learning contexts in the VITAL project

Ilja Cornelisz, Vrije Universiteit Amsterdam
Student Engagement with Computerized Practicing: Ability, Task Value and Difficulty Perception

Data and Theory 1 - Representing digital learning interactions

Room: A2.10   Chair: Marc Esteve del Valle

Regina Motz, Universidad de la República
Detection of Interactions that Impact Learning

Dai Griffiths, University of Bolton
Development of a VLE Recipe for xAPI: process and implications

Marc Esteve del Valle, University of Groningen
Developing Learning Analytics Methods on Reddit

13.00-14.00 Lunch

Location: de Brug [the Bridge]

14.00-15.00 Panel - **Learning analytics policies**

Room: A1.03 Chair: Hendrik Drachsler

Panelists:
Peter van Baalen, Anne Boyer, Jocelyn Manderveld

15.00-15.30 COFFEE BREAK

15.30-17.00 Workshops - LA policy challenges

Workshop 1 - Changing business model of education

Room: A1.03   Moderator: Anwar Osseyran

Issue 1: Innovating the business model of data-driven higher education
Issue 2: Affordable and sustainable LA services
Issue 3: Formalised decision making in the era of data-driven education

Workshop 2 - Presenting learning analytics to stakeholders of education

Room: A2.09   Moderator: Alan Berg

Issue 1: Support for LA Evangelists
Issue 2: Cooperation across competing organisations
Issue 3: Learning from LA failures

Workshop 3 - Learning analytics and organisational culture

Room: A2.10   Moderator: Stefan Mol

Issue 1: Train the trainers
Issue 2: Implementation of personalized education on a larger scale
Issue 3: Privacy and ethics issues related to educational data

19.30-22.30 CONFERENCE DINNER
KIT | Meetings & Events, Restaurant De Tropen

Location: Mauritskade 63,1092 AD Amsterdam

October 27, 2017

10.00-11.00 Keynote - Sanna Järvelä
Room: A1.02

11.00-11.30 COFFEE BREAK

11.30-13.00 Research Presentations

Theory and Methods 2 - Matching education, goals and labour market demand
Room: A1.02 Chair: Scott Harrison

Guanliang Chen, TU Delft, ICSI, UC Berkeley
Buying Time: Enabling Learners to become Earners with a Real-World Paid Task Recommender System

Job Hudig, Rotterdam School of Management, Erasmus University
Long-term effects of study-choice meetings, online personal goal-setting, and an academic stretch goal on student performance

Scott Harrison, University of Siegen, Institute of Knowledge Based Systems & Knowledge Management
“Why do I need This?” - Helping Students Understand Market Demand for Individual Skills

Methods and Data 2 - Self-regulated and adaptive learning

Room: A2.07   Chair: Nicolette van Halem

Christian Weber, University of Siegen

Analysing adaptive learning platforms utilizing domain ontologies: Searching for analytical implications

Vladimer Kobayashi, University of Amsterdam

Investigating the relationships among self-regulation, approach to learning, goal orientation, LMS activity and academic performance.

Nicolette van Halem, Vrije Universiteit Amsterdam

The effect of adaptive practicing on the relation between students' summative grades and following learning activity

Data and Theory 2 - Social comparison feedback and motivation

Room: A2.10   Chair: Chris van Klaveren

Daniel Davis, TU Delft

Follow the Successful Crowd: Raising MOOC Completion Rates through Social Comparison at Scale

David Stap, University of Amsterdam

Gamification in interactive learning environments

Chris van Klaveren, Vrije Universiteit Amsterdam

The Higher Education Enrollment Decision: Bayesian Learners versus Bad News Ignorers

13.00-14.00 LUNCH

14.00-15.00 Panel - **Cooperation between research and industry**
Room: A1.02 Chair: Anwar Osseyran

Panelists:
Stefan Mol, Dai Griffiths, Justian Knobbout, Ian Dolphin

15.00-15.30 **COFFEE BREAK**

15.30-17.00 Sessions - LA Applications in Education

Session 1 - Issues of privacy and ethics
Room: A1.02 Chair: Gábor Kismihók / Stefan Mol

Panelists:
Niall Sclater, Fay Kartner, Nynke de Boer

Session 2 - Early warning systems and predicting student success or failure
Room: A2.07 Chair: Chris van Klaveren / Ilja Cornelisz

Theo Bakker, Vrije Universiteit Amsterdam
Sense Making of Student Analytics, Development and Application of an Early Warning Model to prevent Bachelor Dropout

Lee O'Farrell, National Forum for the Enhancement of Teaching and Learning in Higher Education
DESSI – Ireland's Data-Enabled Student Success Initiative

Jan Hellings, Amsterdam University of applied sciences
The effect of a learning analytics dashboard on the passing rate of a programming course. A randomized controlled experiment

Session 3 - Learning Analytics infrastructure and dashboard development

Room: A2.10   Chair: Alan Berg

Justian Knobbout, HU University of Applied Sciences
Designing a Learning Analytics Capability Model

Ian Dolphin, Apereo Foundation
The Apereo Learning Analytics Initiative. How innovation, community building and 100% Open works on the Global Stage

Nils Siemens, Amsterdam University of Applied Sciences (AUAS)
Learning analytics and lecturers knowledge on learning analytics dashboards

END OF CONFERENCE

ABSTRACTS - OCTOBER 26, 2017

Keynote

Dragan Gasevic
IECS, University of Edinburgh

The field of learning analytics is established with the promise for the education sector to embrace the use of data for decision making. There are many examples of successful use of learning analytics to enhance student experience, increase learning outcomes, and optimize learning environments. Despite much interest in learning analytics, many higher education institutions are still looking for effective ways that can enable systemic uptake. The talk will first describe some selected examples of the successful use of learning analytics in higher education. Key challenges identified to affect implementation of learning analytics will then be discussed. This will be followed with an overview of an approach to the development of institutional policy and strategy for the learning analytics implementation in higher education. The talk will be based on the findings of several international studies and will critically interrogate the role of institutional and cultural differences.

Using Educational Design Research to Develop Actionable Analytics to Support First Year Students

Nynke Bos[1], Maartje Van den Bogaard[1] & Tinne DeLaat[2]
[1]*University of Leiden,*
[2]*University of Leuven*

Purpose: There is increasing attention for the challenges new students face in their first year in university (Mack, 2010). Despite much effort, some groups of students are consistently more likely to struggle to cope with the transition between secondary and higher education (Leese, 2010). Students need to acclimatize to a new environment with new expectations, new information and students need to develop skills to cope with the demands of new educational environment. The first year of college is arguably the most critical regarding the retention of students into subsequent years of study (Arnold & Pistilli, 2012). It is therefore important to guide these students successfully through their first year in university (Charleer, Moere, Klerkx, Verbert, & De Laet, 2017).

In recent years, some universities started to experiment with online applications for student support based on digital footprints students leave in university digital systems (see e.g. Arnold & Pistilli, 2012; Brooks, Greer, & Gutwin, 2014). Many of these experiments focus on predicting student retention on a course level, however, the challenge remains to translate these predictions into actionable results, better understanding of the learning process in question and subsequently design of actions to improve student learning (Lodge,

Alhadad, Lewis, & Gašević, 2017). A theory led design has the potential to yield innovation (Kelly, Thompson, & Yeoman, 2015), opposed to the current atheoretical approach based on the need to leverage the learning data available (Jivet, Scheffel, Drachsler, & Specht, 2017). However, a theory-led approach has the pitfall to ignore contextual factors and impair actionable results.
A research method considering these contextual factors to develop actionable research-based solutions for complex problems in educational practice is educational design research (EDR). Current research explores the added value of an EDR approach to develop learning analytics solution to support first year students into their transition into higher education.

Design: EDR is a research design appropriate to develop research-based solutions to complex problems in educational practice or to develop or validate theories about learning processes, learning environments and the like (Plomp & Nieveen, 2013). The EDR approach consists of preliminary research, such as a needs and context analysis, the development phase consisting of iterative phases aimed at improving and refining the design and the assessment phase to determine if the design meets the specifications.

During the preliminary research phase four sub-studies were undertaken. Two literature reviews were conducted, one regarding student transitions into higher education and one regarding the use of learning analytics to support student retention and success. A contextual sub-study consisting of interview with students and student counsellors and sub-study consisting of contextual analysis based on an institutional baseline measurement of challenges and interventions in the transition into higher education. The preliminary phase resulted in recommendations for using learning analytics to

support student transitions into higher education. One of the main challenges identified for students in their first year, is their lack of reflection on academic grades, and planning after students receive these grades the end of an examination period.
The developmental phase consisted of iterative cycles of designing and using the intervention. The objective of the intervention is to inform and support users with issues that were identified in the preliminary research by means of data analysis techniques to leverage human judgments (Siemens & Baker, 2012). During this cycle, short pilots were evaluated with the student counsellors (observations and surveys) and the students (survey). The iterations and the results of this phase have been extensively covered by Charleer et al. (2017).
The third phase of the EDR consist of reflection to produce 'design principles' and enhance solution implementation. One of the main conclusions of this cycle is that student reflection is not triggered by means of the dashboard itself; analytics becomes actionable when students are offered explicit guidance for awareness purposes and reflections processes to occur (see also Jivet et al., 2017; Charleer et al., 2017). In this case, the student counsellors offered this explicit guidance.

Results: Current research explored the added value of an EDR as a method to develop a learning analytics solution to support first year students with their transition into higher education. It described the iterative process of designing research-based solutions. This resulted in a dashboard to facilitate communication between student counsellors and students by visualising grade data that is commonly available in any institution (Charleer et al., 2017).

Using EDR as a method for designing learning analytics solutions enables researchers to gain fundamental understanding of the underlying processes, goals, context, and even constraints in implementing analytics interventions. Current research shows three distinct processes which will benefit learning analytics. First, the focus on real contexts which is different for experiments. For example, within an EDR approach confounding variables are not controlled for, but are essential for the information they provide ensuring the scalability across contexts and domains. Second, EDR does not simply connect theory and research, but uses theories to build their own theoretical framework for the study (Jen, Moon, & Samarapungavan, 2015). Third, the collaboration between researchers and end-users will assure the uptake of the innovation. In the current research the dashboard will be available institutional wide.

Implications: This research stresses the importance of contextual conditions when designing analytics solution. The recent drive towards theory led learning analytics intervention promote the use of data to assess the effectiveness of educational practices and resources (see e.g. Jivet et al., 2017; Wise & Shaffer, 2015). Critics of these trends argue that education is highly context- specific and practitioner-dependent (McKenney, & Mor, 2015). Current research shows that both are equally important to design actionable analytics, not only the theory determines the intervention, but absolutely key was the importance of contextualising the learning, teaching and counselling in the selected area. This contextualisation assured uptake within the organisation and moreover, an institutional wide adaption of the dashboard.

Acknowledgments: The research leading to these results has received funding from the European Community's Erasmus+ programme, Key Action 2 Strategic Partnerships, of the European Union under grant agreement 2015-1-UK01-KA203-013767 ABLE project.

Resources:

1. Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 267-270). ACM.
2. Brooks, C., Greer, J., & Gutwin, C. (2014). The data-assisted approach to building intelligent technology-enhanced learning environments. In *Learning Analytics* (pp. 123-156). Springer New York.
3. Charleer, S., Moere, A. V., Klerkx, J., Verbert, K., & De Laet, T. (2017). Learning Analytics Dashboards to Support Adviser-Student Dialogue. *IEEE Transactions on Learning Technologies*.
4. Jen, E., Moon, S., & Samarapungavan, A. (2015). Using design-based research in gifted education. *Gifted Child Quarterly*, *59*(3), 190-200.
5. Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough. Pitfalls of learning analytics dashboards in the educational practice. In *European Conference on Technology Enhanced Learning* (pre-print). Springer International Publishing.
6. Kelly, N., Thompson, K., & Yeoman, P. (2015). Theory-led design of instruments and representations in learning analytics: Developing a novel tool for orchestration of online collaborative learning. *Journal of Learning Analytics*, *2*(2), 14-43.
7. Leese, M. (2010). Bridging the gap: supporting student transitions into higher education. *Journal of further and Higher Education*, *34*(2), 239-251.
8. Lodge, J. M., Alhadad, S. S., Lewis, M. J., & Gašević, D. Inferring Learning from Big Data: The Importance of a Transdisciplinary and Multidimensional Approach. *Technology, Knowledge and Learning*, 1-16.

9. Mack, C. E. (2010). Theoretical perspectives on orientation. In J. A. Ward-Roof (Ed.),
10. *Designing successful transitions: A guide for orienting students to college* (Monograph No. 13, 3rd ed., pp. 3-10). Columbia, SC: University of South Carolina, National Resource Center for The First-Year Experience and Students in Transition.
11. McKenney, S., & Mor, Y. (2015). Supporting teachers in data-informed educational design. *British journal of educational technology*, *46*(2), 265-279.
12. Plomp, T., & Nieveen, N. (2007, November). An introduction to educational design research. In *Proceedings of the Seminar Conducted at the East China Normal University [Z]. Shanghai: SLO-Netherlands Institute for Curriculum Development*.
13. Siemens, G., & Baker, R. S. (2012, April). 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.
14. Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, *2*(2), 5-13

SURF Learning Analytics Experiment: Hands-on experience for Dutch higher education

Jocelyn Manderveld and Herman van Dompseler,
SURFnet

Abstract: In 2016 SURFnet started the Learning Analytics Experiment for Dutch institutes for higher education to gain hands-on experience with learning analytics. With this experiment, SURFnet demonstrates the possibilities of learning analytics in education. By carrying out this experiment, educational institutions can answer the following questions: Is learning analytics really so complicated? How does learning analytics fit into an educational infrastructure? How do you collect data? How do you visualise data? In this paper we present the set-up of the Learning Analytics Experiment, the learning analytics architecture and infrastructure used and the institutes who participate in the experiment as well as the first results of the experiment.

Should students have a right to analytics?

Alan Berg, Sjoukje Kerman
Central Services (ICTS), University of Amsterdam

Purpose: The purpose of the presentation is to move away from a focus on barriers to deploying Learning to the benefit of services for students and teachers and their rights around the quality of analytical support. Learning Analytics (LA) is a new field of study, only starting to exist as a unique study since around the early to mid 2010's (Siemens, 2013). The evidence to the effectiveness of LA from the field is inconsistent (Ferguson & Clow, 2017). In addition, full bloodied deployments involve technical conversations around the plumbing, anonymizing and synthesizing of data (Khalil & Ebner, 2016; Berg, Mol, Kismihók & Sclater, 2016) as well as far more importantly teaching approaches (Mor & Wasson, 2015). In this presentation, the authors will describe the barriers to delivering a University wide, consistent and layered system for the aggressive and incrementally improving deployment of Learning Analytics. We will not dwell on the well-known conversations around ethics and privacy (Drachsler & Greller, 2016) or standardization of evaluations of tools (Scheffel et al, 2014) or the details of specific projects (Brouwer et al, 2016). However, we will seek flip the presentation and generate a debate with the audience discussing the following question: "*Should students have a right to analytics?*".

Design: The UvAInform program at the University of Amsterdam consisted of 7 pilots with a number using common and architecture known as a Learning Record Store based on the xAPI standard for capturing online student activity streams (Berg et al, 2016), which is an approach JISC currently apply as part of a National Infrastructure for LA by (Sclater, Berg & Webb, 2015). Although individual pilots within the program were successful from a research and experience building perspective the UvAInform program did not motivate further investment in a scaled-up program for University wide services.

Results: The UvAInform program confirmed that it is difficult for the University of Amsterdam to make top down in combination with bottom up decisions on the deployment of highly technical, difficult to describe yet potentially paradigm changing data driven methodologies. We intend the presentation to generate a further conversation within the communities attending.

Implications: A project is in progress to discuss and inform and listen to decision makers and provide training on the core themes of Learning Analytics. Through this approach we hope to successfully navigate and negotiate and report back on the correct vector for the introduction of new approaches to supporting the University holistically.

Acknowledgments: The authors wish to acknowledge the tireless hard work of those who had labored within the UvAInform project at the University of Amsterdam.

Resources:

1. Berg, A., Scheffel, M., Drachsler, H., Ternier, S., & Specht, M. (2016). Dutch Cooking with xAPI Recipes: The Good, the Bad, and the Consistent. In 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT) (pp. 234–236). IEEE. http://doi.org/10.1109/ICALT.2016.48
2. Berg, A. M., Mol, S. T., Kismihók, G., & Sclater, N. (2016). The role of a reference synthetic data generator within the field of learning analytics. Journal of Learning Analytics, 3, 107–128. http://doi.org/10.18608/jla.2016.31.7

3. Brouwer, N., Bredeweg, B., Latour, S., Berg, A., & van der Huizen, G. (2016). Learning Analytics Pilot with Coach2 - Searching for Effective Mirroring (pp. 363–369). http://doi.org/10.1007/978-3-319-45153-4_28
4. Drachsler, H., & Greller, W. (2016). Privacy and analytics. In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16 (pp. 89–98). New York, New York, USA: ACM Press. http://doi.org/10.1145/2883851.2883893
5. Ferguson, R., & Clow, D. (2017). Where is the evidence? In Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17 (pp. 56–65). New York, New York, USA: ACM Press. http://doi.org/10.1145/3027385.3027396
6. Khalil, M., & Ebner, M. (2016). De-identification in learning analytics. Journal of Learning Analytics, 3(1), 129–138. http://doi.org/10.18608/jla.2016.31.8
7. Mor, Y., Ferguson, R., & Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action. British Journal of Educational Technology, 46(2), 221–229. http://doi.org/10.1111/bjet.12273
8. Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality Indicators for Learning Analytics. Journal of Educational Technology & Society, 17(4), 124–140. http://doi.org/10.1145/2567574.2567593
9. Sclater, N., Berg, A., & Webb, M. (2015). Developing an open architecture for learning analytics. Proceedings of the EUNIS 2015 Congress. http://doi.org/ISSN: 2409-1340
10. Siemens, G. (2013). Learning Analytics. American Behavioral Scientist, 57(10), 1380–1400. http://doi.org/10.1177/0002764213498851

Mastery vs. Avoidance? Impact of grade, sender, comparative information and message style on student affect and academic resilience

Joel A. Howell, Lynne D. Roberts, & Vincent O. Mancini Curtin
University, School of Psychology and Speech Pathology

Purpose: Learning analytics enable automated feedback to students through alerts. However, there is an underlying assumption that simply providing analytics to the student will be sufficient to improve use and self-regulated learning. To date, research exploring student reactions to learning analytics feedback has been limited and largely a theoretical. Working within a framework of supporting learner's agentic engagement with feedback (Winstone, Nash, Parker & Rowntree, 2016) the aim of the present research is to explore student reactions to possible learning analytics messages (alerts).

Design: The present research uses a between-within subjects experimental design. We examined whether varying feedback on alerts for hypothetical assessments based on grade (High Distinction, Pass, and Fail), sender (course coordinator versus automated message), provision of comparative peer achievement, and message style (supportive versus factual) resulted in differences in student affect and academic resilience.

Results: Three hundred and twenty undergraduate students (*Mage* = 22.36 years, *SD* = 6.55 years; 72 males, 245 females, 3 alternative genders) completed an online survey with random allocation to experimental conditions. Multivariate analyses of variance indicated significant differences in affect and academic resilience between grade levels (large effects). Within Pass and Fail grade levels, but not within High Distinction grade level, some smaller effects were observed for comparative peer achievement, message style, and sender.

Implications:The present research has implications for how feedback through learning analytic alerts can best be constructed within each grade level to enhance learner affect and experiences. However, it appears that the key factor as to how students will respond to learning analytics feedback is less about the specific construction of the message than the grade received.

Acknowledgments: The present project was funded by a Curtin Innovation grant

Resources:

1. Roberts, L. D., Howell, J. A., Seaman, K. & Gibson, D. C. (2016). Student attitudes toward learning analytics in higher education: "The fitbit version of the learning world". *Frontiers in Psychology* 7 (DEC): 1-11. doi:10.3389/fpsyg.2016.01959
2. Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a customizable dashboard:personalized learning analytics dashboards in higher education. doi:10.1007/s10758- 017-9316-1
3. Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2016). Supporting learners agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*, 17-37. doi:10.1080/00461520.2016.1207538

Analyzing and visualizing learner behavior with learning analytics in language and mathematics learning contexts in VITAL project

Anouk Gelan, [1] Niels Martin, [1] Martine Verjans, [1] Natasa Brouwer, [2] André Heck, [2] Gerrit Oomens, [2] Victoria Orange-Sibra, [3] Micheal Thomas[3] Ben Betts, [4] Emma Sephton

[1]*Universiteit Hasselt,*

[2]*University of Amsterdam,*

[3]*University of Central Lancashire,*

[4]*HT2*

Purpose: The field of Learning Analytics (LA) opens new possibilities for researching how students learn online, based on the systematic collection and analysis of data about their learning interactions with a variety of online learning environments. The Erasmus+ VITAL project (Visualisation Tools and Analytics to monitor Online Language Learning & Teaching, 2015-2017) aimed to explore these possibilities by implementing LA in 4 different blended or distance learning contexts in 3 European universities. A multidisciplinary team was put together using statistical and process mining techniques to identify learning patterns and learner profiles, to investigate how LA can contribute to better learning design and to analyze whether indicators of success or failure could be discovered. By presenting the results to the students and instructors in the form of learning dashboards visualizing progress and performance, the project aimed to explore how to support students in their autonomous learning process and stimulate their self- reflection and how to allow instructors to monitor their students' progress and struggles so as to adapt their teaching accordingly.

Design: In a first phase a context-specific tracking design was created. Indeed, to deliver useful feedback to students, instructors but also course designers, it is crucial that the implementation of LA is rooted in the pedagogical design of the learning context under analysis. To transcend the local learning contexts of the project a LA design was implemented based on the e-learning specification 'Experience API' (xAPI). This technical specification allows applications to dynamically track, store and share data about learners in their context building on a standardized tracking vocabulary and APIs for learning applications and reporting tools to communicate and exchange data. An open xAPI model describing common tracking vocabularies can allow to capture, analyze and share standardised datasets and answer various online learning research questions. After implementing xAPI tracking in the universities' in-house language or maths learning environments (UHasselt, UvA) or LMS (UCLan), a pilot phase allowed the data analysis team to select and test existing process mining algorithms on different test datasets. The data were collected in the project's central Learning Record Store, more specifically the open source Learning Locker by partner HT2Labs. The pilot phase allowed us to refine the data collection process before collecting, during the main data analysis phase, 4 datasets of 285 UHasselt students, 224 UvA students + 254 UCLan students of semester 1 of 2016-2017.

Results: A descriptive and exploratory research design allowed the different university teams to validate course design hypotheses based on the observed uses of online contents and functionalities. Process mining discovery was used to identify most frequent learning paths throughout the learning environment and a cluster analysis was carried out to profile the learners based on their online behavior.

Based on the research findings, learning dashboards for students and instructors were developed using open source D3.js. For an optimal pedagogical interpretation of the dashboards, local course, student and contents metadata were linked to the xAPI performance data.

After the courses were finished, the learning analytics data collected during courses were visualized on the dashboards. The experiences with these dashboards were evaluated in the 3 universities by students providing them only own data on the personal dashboard and by the instructors providing them only own students' data on the course dashboard.

Implications: The standardized xAPI data format allowed us to build common progress and performance visualization tools while considering the specific learning contexts by selecting the pedagogical indicators considered most relevant for each use case. The LA dashboard design was developed such as to generate future data live from the Learning Record Store, to present these to new student cohorts and to further refine research findings. The xAPI tracking recipes, used open technologies, process mining algorithms, reports, dashboard tools code, dashboard tool user guidelines and recommendations used in the project will be put at free disposal under open licenses.

Acknowledgements: The VITAL project has been funded with support from the European Commission (Project number: 2015-BE02-KA203-012317). The information in this research reflects the views only of the authors, and the Commission cannot be held

responsible for any use which may be made of the information contained therein.

Resources:

1. Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research, 65*(3), 245-281. doi:10.3102/00346543065003245
2. Del Blanco, A., Serrano, A., Freire, M., Martinez-Ortiz, I., & Fernandez, M., B.. (2013, 13-15 March 2013). *E-Learning standards and learning analytics. Can data collection be improved by using standard data models?* Paper presented at the Global Engineering Education Conference (EDUCON), 2013 IEEE.
3. Elias, T. (2011). Learning Analytics: Definitions, Processes and Potential. *retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf, Febr. 2017*.
4. Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *Int. J. Technol. Enhanc. Learn., 4*(5/6), 304-317. doi:10.1504/ijtel.2012.051816
5. Fischer, R. (2007). How do we Know what Students are Actually Doing? Monitoring Students' Behavior in CALL. *Computer Assisted Language Learning, 20*(5), 409- 442. doi:10.1080/09588220701746013
6. Fischer, R. (2012). *Diversity in learner usage patterns Computer-Assisted Language Learning*: Cambridge University Press.
7. Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71. doi:10.1007/s11528-014-0822-x
8. Govaerts, S., Verbert, K., Klerkx, J., & Duval, E. (2010). *Visualizing activities for self- reflection and awareness.* Paper presented at the 9th International Conference on

Web based Learning, Shanghai University, China.
9. Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society, 15*(3), 42- 57.
10. Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *Educause Review, 46*(5), 31-40.
11. Long, P., Siemens, G., Conole, G., & Gasevic, D. (2011). *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK11).* Paper presented at the International Conference on Learning Analytics and Knowledge (LAK11), Banff, AB, Canada.
12. Rogers, T., Dawson, S., & Gasevic, D. (2016). Learning Analytics and the Imperative for Theory Driven Research *The SAGE Handbook of E-learning Research, 2nd edition* (pp. 232-250): SAGE Publications Ltd.
13. Sclater, N. (2016). Developing a Code of Practice for Learning Analytics. *Journal of Learning Analytics, 3*(1), 16-42.
14. Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510-1529.
15. Verbert, K., Govaerts, S., Duval, E., Santos, J., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499-1514. doi:10.1007/s00779-013-0751-2
16. Zinn, C., & Scheuer, O. (2006). Getting to Know Your Student in Distance Learning Contexts. In W. Nejdl & K. Tochtermann (Eds.), *Innovative Approaches for Learning and Knowledge Sharing: First European Conference on Technology Enhanced Learning, EC-TEL 2006 Crete, Greece, October 1-4, 2006 Proceedings* (pp. 437-451). Berlin, Heidelberg: Springer Berlin Heidelberg.

Student Engagement with Computerized Practicing: Ability, Task Value and Difficulty Perceptions

Ilja Cornelisz & Chris van Klaveren

Amsterdam Center for Learning Analytics, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam

Purpose: "When digital tools support students' engagement with challenging material, thus extending learning time and practice, or help students to assume control over the learning situation, by individualizing the pace with which new material is introduced or by providing immediate feedback, students probably learn more. (OECD, p. 166)"

In this study, we aim to find an answer to how task perceptions regarding low-stakes computerized practicing relate to observed patterns of student engagement and performance for students of different ability. Furthermore, we evaluate the impact of personalization on task perceptions and how differences in perceptions and effort evolve over time. Both aspects (i.e. interlinkages between task perceptions and learner readiness and effects of personalization across a heterogeneous group of learners) are currently unresolved issues in the empirical literature (Tomlinson, 2003). To our knowledge, this study is the first to evaluate the relationships between the effort exerted in computerized practicing, student ability and perceptions related to task value and task difficulty constructs.

Design: This research focuses on 455 secondary school students collected throughout the school year 2014-2015. During this period, the academic performance and computerized practicing intensity of students was monitored for 4 different subjects (Dutch, Biology, Economics and History). Afterwards, students were asked about their experiences with computerized practicing, making it possible to link observed differences in practice intensity and student ability to questionnaire items related to constructs of task difficulty and task value.

Both a personalized and non-personalized version of computerized practicing are evaluated in this study and students were randomly assigned to both practicing treatments *within classes*. When students practice with the personalized practicing program relative practicing performance, knowledge type, difficulty level, and mastery learning are taken into account. When the practicing process is not personalized, a predetermined sequence of exercises is offered, which, at least in theory, is representative for the upcoming summative test.

Results & Implications: The results for perceived task interest and usefulness point out that both have the potential to promote student engagement, albeit in different ways and only when students work in a personalized practicing environment. Whether the task is perceived as interesting is related to a higher propensity to practice when a student is assigned to the personalized condition, but conditional on this result, no differences in practice intensity are observed. When the task is considered to be useful, the potential to improve student engagement in the personalized condition is not driven by a higher likelihood to practice in a given session, but if students take up this opportunity, they will do so for a longer period of time.

In evaluating the implications of interest and perceived usefulness, only usefulness corresponds to higher levels of practice intensity and only for students assigned to the personalized condition. Again, none of these differences are mirrored by differential patterns in performance on summative tests. Exploring how practice intensity evolved over time furthermore reveals that students first need to gain experience with the task, before differences in perceived attitudes emerge and can be correlated to corresponding differences in practice intensity. When disaggregated by ability, results regarding task perceptions are markedly different for the two versions of computerized practicing. In the non-personalized condition, students with relatively lower pre-scores value practicing as useful and as too difficult, while in the personalized condition students with relatively higher pre-scores value this process as useful. One plausible interpretation for this result is that students of relatively lower ability consider practicing useful if it closely resembles the summative test for which they are preparing, whereas higher ability students attribute more usefulness value when practicing is relatively challenging and more directly addressing their personal learning needs.

Future Research: For future research, it is furthermore important to acknowledge that one explanation for the suboptimal results of current computerized personalized practicing tools may well be that adaptive processes are generally offered in a one-size-fits-all approach, as was the case in this study. The results presented in this paper with respect to ability suggest that students would assign more task value to computerized practicing if the process would properly take into account the heterogeneity of the student population. It can be argued that there are infinite ways to personalize the learning process and optimal adaptation of the content offered to a heterogeneous group of learners cannot be realized using only a single algorithm. Optimally, a system continuously evaluates for each student the algorithm which best accommodates the individual learning needs and preferences. As a result, it may well be that multiple algorithms are to be developed and implemented in order to continuously adapt the practicing process to individual needs.

Detection of Interactions that Impact Learning

Regina Motz[1], Ofelia Cervantes[2], Blanca Viera[3]

[1]*Universidad de la República, Uruguay*

[2]*Universidad de las Américas Puebla, México*

[3]*Administración Nacional de Educación Pública/CERP del Suroeste, Uruguay*

Purpose: The *Ceibal Plan* is a socio-educational project of Uruguay, created by decree of April 18, 2007, "to carry out studies, assessments, and actions necessary to provide a laptop to each school-age child and to each public school teacher, as well as to train teachers in the use of this tool, and to promote the elaboration of educational proposals in line with them" [1]. The Ceibal Plan (the One Child Laptop program implemented in Uruguay) seeks to promote digital inclusion to reduce the digital gap, both in relation to other countries and among the citizens of Uruguay. Moreover, it provides educational support services, using information technologies, to all students in public education at the primary and secondary level in Uruguay. It emphasizes the use of two platforms: CREA 2 (Learning Management System offered by Schoology) and PAM (Adaptive Platform for Mathematics offered by Bettermarks). CREA 2 reporting in May 2015 a daily activity of 85,000 users per hour and PAM a daily activity of 95,000 active users per day in 2014. Despite this numbers, the report entitled "Deepening the effects of Ceibal Plan" [2], conducted by professionals of the Institute of Economics from Universidad de la República and funded by the plan itself and the National Public Education Administration (ANEP), states that the distribution of laptops has not generated an improvement in the academic performance of students. This report also points out that it is considered "primordial" the "develop strategies aimed at promote teacher empowerment and the creation of collective capacities to focus on teaching and learning through access and innovative use of technology."

Teachers with initiative and concerned about the improvement in the academic performance of their students have explored the use of formal social networks existing in online communities (such as CREA and PAM) and informal social networks (such as Facebook), as spaces to stimulate the search and construction of knowledge through interaction between students, as well as between students and the teacher.

We work in a project that aims to the development of a software platform that allows teachers to visualize patterns of interaction and relate them to learning levels in a student- centered way regardless of the spaces (Facebook, CREA, PAM) where the data is generated. Patterns of these interactions are critical information that supports teachers in making strategic decisions for improving the academic performance of students. However, we found that primary and secondary level teachers are not yet in a good relation with learning analytics technologies. In our presentation, we will show the method used to attract teachers towards the use of software for learning analytics.

Our proposal is to present the progress of the project we are conducting as an interdisciplinary work among computer scientists, educators and psychologists to develop a tool that assists teachers to discover interactions that occur in social networks and analyze how they impact on learning [3]. The approach builds an integrated profile of the student, incorporating their demographic, educational and social characteristics. Applying information retrieval techniques, we capture student's facets from institutional and non-institutional social networks. The data is provided both from their activity in the Learning Management Systems (CREA2 and PAM)) as well as in from informal social networks such as Facebook, Google+ or Twitter. These non-institutional social networks (self- regulated social network spaces) are relevant spaces because their students behave in a different way than in the LMS, among their colleagues. The types of interactions analyzed are student-material, student-students and student-teachers. We apply traditional learning social analytics models increasing with approaches that contemplate the semantic nature of interactions to capture interactions quality, these may range from positive to a negative interaction.

A pilot study is conducted to validate the benefits of using the platform to support the teacher in detecting cases requiring specialized care, such as isolation (possible depression), bullying, strategic communication agents, etc. The pilot scenario is within the framework of teacher training groups. This allows us to work with students of legal age and who are also proactive in the use of technology, which mitigates the risk of not having enough interactions in the platforms to study. On the other hand, it is one of the possible ways produce changes as they are generated from the new generations of teachers. Further, we propose a rush formation plan for all teachers not familiar enough with learning analytics technologies.

We seek to provide the teacher with timely access to relevant information in their students learning process to assist them in designing inclusive education strategies. For this, it is important to give teachers a work environment with data that allows them to take a proactive nature, and that offers relevant contents according to the observed activity and individual interests specified through student profiles. In this sense, the data to be analyzed is not only quantitative but rather qualitative in nature.

Data provided by Ceibal Plan on the use of institutional resources by students, and incremented by our project, included, but are not limited to:

- Frequency of connectivity, time and place from which students are connected,
- Type of resources and frequency of use thereof, for each student,
- Frequency of access and use of social networks, and quality positive or negative in their interactions.
- Performance in school activities.
- Timely attendance to face-to-face classes
- Attitude in face-to-face classes

We present the analysis of social networks and from data generated in face-to-face class focused on each student to enrich his/her profile and identify patterns of interactions that impact on learning. Moreover, the student-centered approach identifies colleagues who support that student's learning. This method has also the potential to help identify groups within the network, which can support learning processes, such as communities and affinity groups.

Resources:

1. Ceibal Plan. http://www.ceibal.edu.uy/.
2. Deepening the effects of Ceibal Plan. Alina Machado. http://www.elpais.com.uy/informacion/plan-ceibal-investigacion-rendimiento- matematicas-lectura.html
3. DIIA-Project: Descubrimiento de Interacciones que Impactan en el Aprendizaje. Funded by ANII-Fundación Ceibal, Uruguay, 2017.

Development of a VLE Recipe for xAPI: process and implications

Dai Griffiths[1], David Sherlock[2], Alan Paull[2]

[1]*University of Bolton, Associate Member of Cetis LLP*

[2]*Cetis LLP*

Purpose: The Jisc Effective Learning Analytics Initiative is “working in collaboration to build a learning analytics service for the sector”, with “over 50 universities and colleges signed up to the initial phases of the implementation” (Jisc, 2017). Cetis LLP was awarded a contract by Jisc to support the development of xAPI recipes for the Initiative. This paper describes the work carried out and its implications.

Design: Data inputs to the Effective Learning Analytics system comes from two sources. Firstly, data is gathered from institutional systems, which maintain records of students’ identity, courses, assessment results, etc. The requirements of the UK Higher Education Statistics Agency (HESA, n.d.) provide some coherence, but there remain inconsistencies which are barriers to a sector wide analytics service. Consequently, Cetis LLP were asked to work on a Universal Data Definition (UDD). Readers interested in this work can consult the Jisc Learning Analytics Unified Data Definitions, currently in version 1.3 (see ‘Resources’ below).

Secondly, data is gathered from the interactions between learners and institutional systems, particularly Moodle and Blackboard, and xAPI is used to ensure that this data can be consumed reliably by the analytics systems. To this end, Cetis LLP has worked with Jisc to define a set of xAPI recipes, which is now available in version 1.0 (see ‘Resources’). Cetis LLP has facilitated dialogue with vendors and education institutions, maintained the Github repository, and resolved issues raised them, with input from Jisc when needed.

Results: Release 1.0 of the VLE recipe, August 2017, consists of a set of platform-independent statement templates that send data to the Jisc Learning Record Warehouse. Full statement examples are included, and the data needed to create the statement is identified. The statement templates are:

- Logged in
- Logged out
- VLE resource viewed
- Assignment graded
- Assignment submitted

‘Forum contribution’ and ‘Library loan’ are scheduled for 1.1. Examples for Blackboard and Moodle are provided. As far as possible all entities are the same across statements. To this end, a common vocabulary was developed, with IRIs and definitions for verbs, activity types, etc, as well as for extensions used in the recipes. A set of common structures represents actors, verbs, objects, contexts and results. Work has also started on recipes for ‘Attendance’ and ‘Mobile App Usage’, with a single statement provided in each recipe.

Implications: When the team has been asked to provide an xAPI statement for a particular purpose, the specification has proved sufficiently powerful and flexible, with clear guidance on how to construct an appropriate statement. We have seen no technical problems to cause us to doubt Ben Betts of HT2 Labs, who asserted that “the adoption rate of xAPI is probably unprecedented in our industry” (Betts, 2017). We also note the excellent work underway in developing the necessary infrastructure, for example the Apereo Learning Analytics Initiative (see resources). Our uncertainties, however, have emerged from engaging with vendors, institutions, and analysts, who have a wide range of priorities and perspectives. It is relatively easy for vendors to generate the xAPI compliant JSON from their applications, and many have done so, but it is more complex to work with stakeholders to establish what this data represents, and how it should be processed. Indeed, the relatively small number of recipes which we have developed in v1.0 hides the richness of the conversations informing the design, as shown by the fact that in the first 12 months of the project the Cetis LLP team resolved 96 issues and made 302 commits on GitHub related to the xAPI work.

We have developed a recipe for use with VLEs, i.e. “a way of expressing how a common type of learning activity could be syntactically represented” (ADL 2016, p.19). We have also provided a vocabulary, which has been the focus for much of the discussion with institutions and vendors. ADL (2016, p.19) associates vocabularies with profiles, rather than recipes, and our experience suggests that the development of effective, shareable vocabularies and profiles will be critical to the further adoption of xAPI. There are, as yet, few profiles and vocabularies available as examples. Moreover, the development of profiles is complex. Firstly, the flexibility of xAPI leads to a temptation to create new statements for every stakeholder request, and to stretch the specification to facilitate analysis. Secondly, Jisc have shown exemplary commitment to working with the community of adopters. Nevertheless, in any product development process, there is limited time to discuss each profile decision with unlimited stakeholders. There is no established method for reconciling the needs stakeholders. We invented the process as we went along, starting in Google Docs, and then moving to GitHub, and felt the need for guidelines.

Some of our stakeholders requested queries for the high-level concept ‘intervene’; others wanted to distinguish between interventions (e.g. automated interventions, email interventions, and face-to-face interventions), and when a student was passive recipient of an activity. In practice, we might expect that many stakeholders would like to query at both levels, requiring nesting. The specification is clear that “A SubStatement MUST NOT contain a SubStatement of its own, i.e., cannot be nested” (ADL, 2012). It is possible to add information to the context property, “such as the instructor for an experience, if this experience happened as part of a team-based Activity, or how an experience fits into some broader activity.” (ADL, 2012). However, this approach would lead to the development of ad hoc ontologies of activities for each profile, which would be hard to inspect or share. ADL recognised this problem in the Companion Specification for xAPI Vocabularies (see resources), recommending a Linked Data representation of the relationship between vocabulary items. At the end of 2016 Cetis LLP recommended this approach for future Jisc work. Many details about how to approach this remained, however, unclear. Since then, ADL and DISC have created a profiles specification to “improve practices for creating Profiles”, making use of Linked Data (ADL, 2017). Our experience indicates that this is a necessary step with the potential to greatly increase adoption of xAPI.

Acknowledgments: The work described in this paper was funded by Jisc. We would also like to acknowledge the support and contributions of the Jisc team throughout the project.

Resources:

1. Jisc Effective Learning Analytics https://www.jisc.ac.uk/rd/projects/effective-learning-analytics
2. Jisc Learning Analytics Unified Data Definitions repository: https://github.com/jiscdev/analytics-udd/
3. Jisc xAPI recipe repository: https://github.com/jiscdev/xapi
4. Apereo Learning Analytics Initiative https://www.apereo.org/communities/learning-analytics-initiative
5. xAPI specification https://github.com/adlnet/xAPI-Spec/blob/master/xAPI-About.md
6. Companion Specification for xAPI Vocabularies, 1.0 (2016) https://www.gitbook.com/book/adl/companion-specification-for-xapi-vocabularies/details
7. xAPI Profiles specification https://github.com/DataInteroperability/xapi-profiles/blob/master/xapi-profiles-structure.md

Reference list

1. ADL (2012) *Experience API*. Available at https://github.com/adlnet/xAPI- Spec/blob/master/xAPI-About.md (accessed August 3rd, 2017).
2. ADL (2016) *Companion Specification for xAPI Vocabularies 1.0.* Available at https://adl.gitbooks.io/companion-specification-for-xapi-vocabularies/content/relationship_between_vocabularies,_profiles,_and_r.html (accessed August 3rd, 2017).
3. ADL (2017) *xAPI Profiles specification.* Available at https://github.com/DataInteroperability/xapi-profiles/blob/master/xapi-profiles-structure.md (accessed August 3rd, 2017).
4. Betts, B. (2017) *xAPI: Where It's Come From, and Where It's Going Next (in 2017).* Available at https://www.ht2labs.com/blog/xapi-where-from-where-next/ (accessed August 3rd, 2017).
5. HESA (n.d.) *Experts in UK higher education data and analysis*. Available at https://www.hesa.ac.uk/about (accessed August 3rd, 2017).
6. Jisc (2017) *Effective Learning Analytics*. Available at https://www.jisc.ac.uk/rd/projects/effective-learning-analytics (accessed August 3rd, 2017)

Developing Learning Analytics Methods on Reddit

Priya Kumar, *Ryerson University*
Anatoliy Gruzd*, Ryerson University*,
Caroline Haythornthwaite, *Syracuse University*
Sarah Gilbert, *University of British Columbia*
Marc Esteve Del Valle, *University of Groningen*
Drew Paulin, *University of California, Berkeley*

Purpose: This presentation introduces the 'learning in the wild' coding schema, an approach developed for learning analytics research and scholars interested in better understanding the different types of discourse, exploratory talk, and conversational dialogue happening on social media. It considers how learner-participants ('Redditors') are leveraging subreddit communities to facilitate self-directed informal learning practices on the Reddit social networking site. Reddit is an online news sharing site that is commonly referred to as 'the front page of the Internet' for the way it presents headlines and how crowd-based online voting raises the profile of news or other information to a front page equivalent. Reddit has become increasingly popular since its launch in 2005, and now maintains a relative stronghold as the go-to, self-organized community site for people interested in current affairs, social commentary and Internet subcultures. The presentation reports on the development of a coding schema for content analysis of informal learning on social media derived by examining the kinds of learning happening on Reddit, and shares results on the kinds and distribution of learning practices found in four 'Ask' subreddit communities ('AskHistorians', 'Ask_Politics', 'askscience', 'AskAcademia'). The research brings attention to the new types of collaborative knowledge, ideas and

resources being shared and supported outside the confines of traditional education and professional environments.

Design: In developing our coding schema, we followed on Ferguson and Buckingham Shum in their work of identifying elements of *exploratory dialogue* in a manner suitable for machine learning [1, 2]. Like Ferguson and her colleagues, we build on Mercer's exploratory talk because it represents the kind of constructive, collaborative interaction that reflects adult, interactive learning and is likely to advance both individual and group knowledge [2, 4, 7]. A focus on exploratory learner dialogue fits well with Reddit because the platform maintains a user- generated participatory online culture through its informal, openly accessible, group-based subreddit communities [8].

The process of developing the coding schema comprised three stages, with Ferguson et al.'s (2013) cue phrase framework comprising seven of the nine categories in Version 1. We used DiscoverText, a cloud-based text-analysis software program [9] that allowed assigning multiple coders to the same dataset. The first cycle of coding was undertaken on a dataset of 1% of 2015 subreddit posts (excluding parent submissions) from each of 'Ask_Politics' (n=189), 'AskAcademia' (n=197) and 'askscience' (n=163). Each sample was coded by three coders. Krippendorf's alpha statistics on intercoder reliability showed a relatively low agreement among coders ('Ask_Politics' 0.16, 'AskAcademia' 0.2 and 'askscience' 0.22). Through a process of iterative refinement, stages 2 and 3 focused on resolving inconsistencies and improving the coding schema.

Version 3 (our final version) of our coding schema is a significant departure from Ferguson et. al (see Table 1). In this third cycle of refinement, we simplified the categories to facilitate coders' use of the codes, standardize multi-coder agreement, and address more specifically the types of exploratory learning dialogue that we were observing on Reddit. Version 3 captures two trends observed in reading Reddit comments: the positive expressions and supportive dialogue and information provision that pull participants toward each other and foster topic-specific discussions, and the more negative exchanges that monitor and sanction behaviour, silence participants, and can stifle online learner dialogue.

Table 1. Coding Schema (final version)

Code	**Definition**	**Linguistic Dialogue Example**
1. Explanation with Disagreement	Expresses a NEGATIVE take on the content of the previous comment by adding new ideas or facts to discussion thread.	‘But’, ‘I disagree’, ‘not sure’, ‘not exactly’ with explanation/ judgment/ reasoning/ etc.
2. Explanation with Agreement	Expresses a POSITIVE take on the content of the previous posts by adding new ideas or facts to discussion thread.	‘Indeed’, ‘also’, ‘I agree’, with explanation/ judgment/ reasoning/ etc.
3. Explanation with Neutral Presentation	Expresses a NEUTRAL explanation/judgment/reasoning/etc. with neither negative nor positive reference to the content of the previous comments, nor necessarily any reference to previous comments.	Comments with non- judgmental language. Advice, brainstorming and first hand experiences are framed neutrally. ‘I can understand’, ‘interesting’, ‘depends on…’ or statement responses.
4. Socializing with Negative Intent	Socializing that expresses negative affect through tone, words, insults, expletives intended as abusive.	‘no’, ‘you’re an idiot’, ‘this has been explained multiple times’

5. Socializing with Positive Intent	Socializing that expresses positive affect tone, words, praise, humor, irony intended in a positive way.	'thanks', 'great feedback', 'you're correct'
6. Informati on Seeking	Comments asking questions or soliciting opinions, resources, etc. ('Does anyone know …?' 'How does this work?'). This does not include questions answered rhetorically within the comment, e.g., if a question is asked and answered.	'First you have to think what happens if …?' and then you can see what happens', 'does anyone know', 'can anyone explain'
7. Providing Resources	Comments that include direct reference to a URL, book, article, etc.; comments that call upon a well-known theory or the name of a well-known figure.	Link to resource copied (book, URL, article, audio/video file). Referencing theory/theorists, scholar or public work (Einstein, Newton, Freud).
8. Subreddit Rules and Norms	Comments on topics such as what is the appropriate sub-reddit for a particular discussion, what language is appropriate to use, how to back up claims by using resources, etc.	'See/don't forget subreddit link', 'this post doesn't belong here', upvote/downvote mentions, acknowledging OP redditors, and bots.

Results: We will show the utility of our coding schema when studying unstructured, informal learning processes through analysis of four 'Ask' subreddit communities. Results of our coding test for Version 3 showed a more acceptable level of agreement (Krippendorf's alpha) between coders: 'Ask_Politics' 0.52, 'AskAcademia' 0.64 and 'askscience' 0.67. In preparation for our validation processes, we also tested the final version of the coding schema with 'AskHistorians' 2015 subreddit sample (n=267) and recorded an alpha of 0.57. While these values are considered of moderate agreement, they are much stronger than in Version 1 of our coding schema (see Design). Three independent coders were then used to test the validity of the schema on a larger, more recent dataset (2016 'AskHistorians' sample) and recorded an alpha of 0.76 (79% intercoder agreement). We regard this alpha level to be acceptable, when considering that we allowed multiple codes (maximum 3) per comment. For exploratory studies like ours, alpha levels between 0.67 and 0.80 are considered reliable enough to draw out and develop cautionary

Conclusions [2, 5, 6]: The results also show that our coding schema can capture subtle nuances in the way people converse across different subreddits (see Table 2). Distribution results from 'AskHistorians' and 'askscience' show that online conversations and social learning processes connect people, Q&A transactional dialogue and external resources. In both cases, we found subreddit community norms to promote civility, collaboration and participatory dialogue, which help encourage self-directed learning practices. The 'ask_Politics' distribution results conversely shows a greater proportion of comments with negative socializing, disagreement and debate which may influence processes of learning (and even unlearning). In contrast to these subject- led subreddits, the professionally-focused 'AskAcademia' subreddit highlights a new range of self-directed learner practices that do not necessarily have a curricula/subject counterpart. Comments in this subreddit were found to be more neutral, supportive, reflective and socially positive; appealing to budding academics by focusing on personal needs.

Table 2. Coding Results*

	ask_Politics	askAcademia	askscience	askHistorians	askHistorians
Year	2015	2015	2015	2015	2016
Sample Size	190	198	164	267	1,227
1.Explanation with Disagreement	91 (48%)	21 (11%)	16 (10%)	34 (13%	71 (6%)
2.Explanation with Agreement	11 (6%)	20 (10%)	10 (6%)	4 (1%)	45 (4%)
3.Explanation with Neutral Presentation	45 (24%)	102 (52%)	100 (61%)	67 (25%)	592 (48%)
4.Socializing with Negative Intent	37 (19%)	5 (3%)	0 (0%)	0 (0%)	4 (0%)
5.Socializing with Positive Intent	2 (1%)	44 (22%)	19 (12%)	31 (12%)	204 (17%)
6.Information Seeking	22 (12%)	13 (7%)	23 (14%)	29 (11%)	274 (22%)
7.Providing Resources	20 (11%)	13 (7%)	33 (20%)	64 (24%)	260 (21%)
8.Subreddit Rules and Norms	3 (2%)	6 (3%)	2 (1%)	0 (0%)	66 (5%)

*Note: For the 2015 'training' datasets, the counts represent an agreement between two or more independent coders. Comments where two or more coders did not agree were not counted or included. For the 2016 validation dataset, the counts represent an agreement between two or more independent coders. Percentages may be higher than 100% when coders have assigned multiple (maximum three) codes per comment.

Implications: The research reasserts the potential of social media sites such as Reddit to support self- motivated learners and sustain communities of practice. In doing so, we highlight different spheres of knowledge, informal learning practices and exploratory dialogue that occur in online settings, outside of traditional classroom environments [3]. We intend to expand this research agenda, first with a larger sample of subreddits, and then across other social media platforms (e.g. Twitter, Facebook, LinkedIn). By detailing our process of coding schema refinement, we invite other scholars to apply the coding schema to their research on informal learning in open, online environments. Upon further validation, we intend to integrate automatic machine learning to our research with the goal of improving models for learning analytics.

Acknowledgments: This work is supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) grant, “Learning Analytics for the Social Media Age”, PIs: Anatoliy Gruzd and Caroline Haythornthwaite.

Resources:

1. Ferguson, R., and S.B. Shum, “Social Learning Analytics: Five Approaches”, in Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM, Vancouver, Canada, 2012, pp. 23-33.
2. Ferguson, R., Z. Wei, Y. He, and S. Buckingham Shum, “An Evaluation of Learning Analytics to Identify Exploratory Dialogue in Online Discussions”, in Proceedings of LAK’13, Leuven, Belgium, 2013, pp.85-93.
3. Gruzd, A., D. Paulin, and C. Haythornthwaite, “Analyzing Social Media and Learning through Content and Social Network Analysis: A Faceted Methodological Approach”, Journal of Learning Analytics, 3(3), 2016, pp.46-71.
4. Haythornthwaite, C., and R. Andrews, E-learning Theory and Practice, Sage, London, 2011.
5. Krippendorff, K., “Reliability in Content Analysis”, Human Communication Research, 30(3), 2004, pp. 411–433.
6. Krippendorff, K., Content Analysis: An Introduction to Its Methodology, Sage, Newbury Park, CA, 1980.

7. Mercer, N., “Sociocultural Discourse Analysis: Analysing Classroom Talk as a Social Mode of Thinking”, Journal of Applied Linguistics, 1(2), 2004, pp.137-168.
8. Moore, C., and L. Chuang, “Redditors Revealed: Motivational Factors of the Reddit Community”, in Proceedings of the 50th Hawaii International Conference on System Sciences, 2017, pp. 2313-2322.
9. Shulman, S., “DiscoverText: Software Training to Unlock the Power of Text”, in Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, ACM, NY USA, 2011, pp. 373-373.

Workshop 1 - Changing business model of education

The explosive rise of higher education (HE) costs and related tuitions, growing dissatisfaction of students about their ROI, and the increasing pressure on institutions and governments to reform higher education systems have undermined traditional HE business models. In fact one may question whether there is a common understanding of HE business models among the various stakeholders of HE such as students, tutors, staff, faculty, boards of management and trustees. Analysing and more importantly modernizing the business model is a major challenge for the HE sector. How to evolve from traditional HE systems with formalized decision making and separated governance towards a more networked data-savvy organization while at the same time improving the outcome? To what extent are these developments confined to HE or do they also have important implications for other sectors in education?

Issue 1: Innovating the business model of data-driven higher education

The existing business models in HE reward spending rather than cost-efficiency because of the significant complexity of cost analysis and uncertainty of potential benefits of cost-cuttings. The philosophy here is that in order to make effective changes, we need to understand how the different activities drive spending and revenues and how they influence learning goals. The approach requires transparency and inter-institutional collaboration to aggregate activity-based cost and results data which reveal competition- and reputation-sensitive patterns. The patterns need to be managed in order to embrace the sector-wide collaboration required for the development of constantly evolving business models.

Issue 2: Affordable and sustainable LA services

Gartner (2016) reports that despite all these benefits and the potential to provide answers to key challenges in education, only a small number of HE institutions engage in institutionally scoped LA. The same study also reveals that LA has nontrivial positive effects on student outcomes, that have the potential to avoid further financial losses of the sector. However, HE organisations still concentrate on institutional (administrational) analytics, rather than focussing on learning and education. This is surprising in light of the fact that besides privacy issues *"Affordability is the biggest barrier to implementing learning analytics. However, the costs of losing/replacing students, as well as the costs of declining government support due to unmet goals, vastly exceed any investment made in learning analytics solutions." (Gartner report, p.1)*

Issue 3: Formalised decision making in the era of data-driven education

LA dashboards are more and more common in teaching and learning and help the work of teachers and students. Dashboards using data from student administrational systems and other financial and administrational (e.g. HR, logistics) systems have been widely used in educational institutions to make decisions. However, success in learning is one of the key output of education, and this type of data and analytics should be represented also in managerial decision support.

Workshop 2 - Presenting learning analytics to stakeholders of education

Isn't it safer to do nothing? How do we present Learning Analytics to stakeholders to optimise the chances of safely navigating this difficult to understand technology-driven paradigm shift when such processes require significant cognitive and financial investments?

Big Business is thriving on Big data. New opportunities are arising; data scientists are as rare as unicorns, and with every opportunity, there are risks such as increasing specific occupational gender imbalances, the desolation of traditional jobs and the replacement of human judgement by AI with their hidden and hard to explain biases. One can make the argument that the educational sector is not adapting fast enough to incorporate the dominance of data and embrace a data-driven mentality as expressed through Academic and Learning Analytics.

Issue 1: The lack of support for Evangelists

There are many LA deployment models, either driven by top-down leadership or bottom-up or a combination. Each model requires a significant degree of evangelism as there is on average a great distance between our, Higher Ed's current status and that of the degree of adoption by data-driven businesses. Key issues in need of being addressed include (but are not limited to) leader/evangelist empowerment, engaging the research community, and leveraging/expanding evidence hubs such as LACE.

Issue 2: The inability to work well together across competing organisations to improve the quality of education for all

We like to think that we are unique, are we just being inefficient? We can argue that the deployment of LA is indeed just a bundle of commonly accepted learning strategies, best practices and design criteria captured with implementation. Reviewing uniqueness, we need to divide that which is exclusive to our organisations and that we could and should we share. Cost-effective, high-quality Education through the promotion of standardisation and shared infrastructure. Striving, struggling for one data governance model connecting nationally scaled systems, creating standard practices, benchmarking predictive models, knowing our algorithmic and teaching biases. Working together well celebrating our uniqueness.

Issue 3: The lack of honesty around failure

Without failure, there is no meaningful learning growth. We have many examples; InBloom, the NHS data release to Google, racist algorithms, fake news, counter-intuitive laws, ageism, stakeholder disconnect, organisational culture. However, failing is not an option and reporting failure as well. I would argue that we have lost the ability to take contrary evidence around failed projects or methods and turn it into constructive improvement. We need to work out how we are going to collect, analyse, advertise and incorporate failures for the betterment of further deployment, categorising and dissecting failure as a community. We will need to expand our definitions and learn 100 percent openly.

Workshop 3 - Learning analytics and organisational culture

Data driven education poses big challenges to stakeholders of education. These challenges are pushing members of such communities into a difficult position, where they need to reconstruct their tasks and activities in order to maintain their function in education. This often means that their own professional identity is at stake (for instance that of teachers). This is the result of 1) a lack of insight in the impact of present educational systems and practices; 2) a lack of resources to provide cost-effective education; 3) a lack of up-to-date learning content and teachers who can competently leverage it to their advantage; and 4) a lack of demonstrated effectiveness of educational programmes and practices. Early adopters of disruptive educational technology are already gearing up and placing the aforementioned issues on the agenda. These organisations have been building their services on educational technology and data for decades, therefore analytics is deeply embedded in their organisational culture. However, these organisations still represent the minority of educational programs and their activities have as of yet not resulted in a major transformation of the educational landscape.

Issue 1: Train the trainers

LA will enable the stakeholders of education to have greater understanding of how students learn and what may be done to enhance their learning. This will result in the dynamic development and evidence based validation of new educational products, where students gain more and more control over their learning process and personalized and adaptive online learning environment. Teachers and trainers need to be prepared for this new world of education and their preparations should be driven not only by their individual motivation to adapt, but also organisations, who need to make a sustained effort to help them to develop and succeed.

Issue 2: Implementation of personalized education on a larger scale

Personalising learning content to individual needs is one of the key assets of Learning Analytics: it allows focussing resources on the students that need those resources the most efficiently and in a timely fashion. Aligning learning content to personal learning goals, skills, and educational performance includes the rethinking of existing grading and performance systems, where for instance tests and assessments should scaffold learning (e.g. by identifying learning needs) rather than evaluate it (formative vs summative assessment).

Issue 3: Privacy and ethics related to educational data

How may we exploit the wealth of information obtained by gathering data with the explicit consent of students and tutors? Coupling data gathered in the class rooms, through sensors, and/or in online courses with other data sources such as social media and student social and educational profiles may be expected to increase the effectiveness of LA, at the risk of simultaneously increasing breaches of privacy and/or ethical standards. All stakeholders, managers, regulators, teachers, students therefore need to be involved and help develop new legislation that does not inhibit but rather facilitates the evolution of LA while at the same time prohibiting misuse.

ABSTRACTS - OCTOBER 27, 2017

Keynote

Sanna Järvelä*,*
LET, University of Oulu

Understanding how people learn is critical for helping them to learn better. A major problem is that the core learning processes – cognition, affect, metacognition – are not visible to the teacher, the learner and the collaborators and, thus, difficult to support by the teacher or technological tools. Self-regulated learning (SRL) theory has helped us to understand the critical processes of learning. Self-regulation is an invisible complex cognitive, motivational, and emotional mental activity that must be acquired by learners and supported by teachers, tools, and environments. While SRL is difficult for individuals, it is even more so when they interact with peers and in teams; co-regulation (CoRL) and socially shared regulation (SSRL) of learning respectively. In my talk, I will demonstrate how we have been using multimodal data (i.e., physiological data, log traces, eye tracking, video-data, contextualized self-reports) to understand a process of learning and socially shared regulation in groups. Since more complex and larger amounts of data are now available than ever in the past, we have wrestled with limitations in our own thinking, as well as with problems related to data handling and analysis and tried to solve those problems in multidisciplinary teams. The methods and tools for learning analytics have helped us trace and model SRL processes and understand learning in authentic learning contexts.

Buying Time: Enabling Learners to become Earners with a Real-World Paid Task Recommender System

Guanliang Chen[1], Dan Davis[1], Markus Krause[2], Claudia Hauff[1], Geert-Jan Houben[1]

[1] *TU Delft Lambda Lab,*

[2] *UC Berkeley ICSI*

Purpose: Massive Open Online Courses (MOOCs) aim to educate the world, especially learners from developing countries. While MOOCs are certainly available to the masses, they are not yet fully accessible. Although all course content is just clicks away, deeply engaging with a MOOC requires a substantial time commitment, which frequently becomes a barrier to success. To mitigate the time required to learn from a MOOC, we here introduce a design that enables learners to earn money by applying what they learn in the course to real-world marketplace tasks. We present a Paid Task Recommender System (Rec-$ys), which automatically recommends course-relevant tasks to learners as drawn from online freelance platforms. Rec-$ys has been deployed into a data analysis MOOC and is currently under evaluation.

Design: We designed and implemented a Paid Task Recommender System (Rec-$ys), which automatically collects course-related tasks from UpWork and recommends them to learners. As depicted in Figure 1, the system structure mainly consists 4 layers:

- MOOC: The MOOC layer serves as the playground for learners to interact with course components as well as Rec-$ys.
- Data layer: This layer is responsible for: 1) keeping track of learners' activities and 2) collecting paid tasks from on- line freelance platforms.
- Analysis layer: This layer aims to analyzing the relevance of tasks for learners based on their interaction with Rec-$ys.
- Intervention layer: To avoid a learner keeps receiving the same task or tens of thousands of learners compete for the same task, this layer dedicates to diversifying tasks recommended to different learners.

Results: Rec-$ys has been deployed in a MOOC which runs from November 22, 2016 to May 23, 2018 in a self-paced mode. Based on findings from [2], we demonstrated that:

- Learners are able to solve real-world freelance tasks in high accuracy and quality
- Real-world freelance tasks are beneficial for improving learners' course engagement.

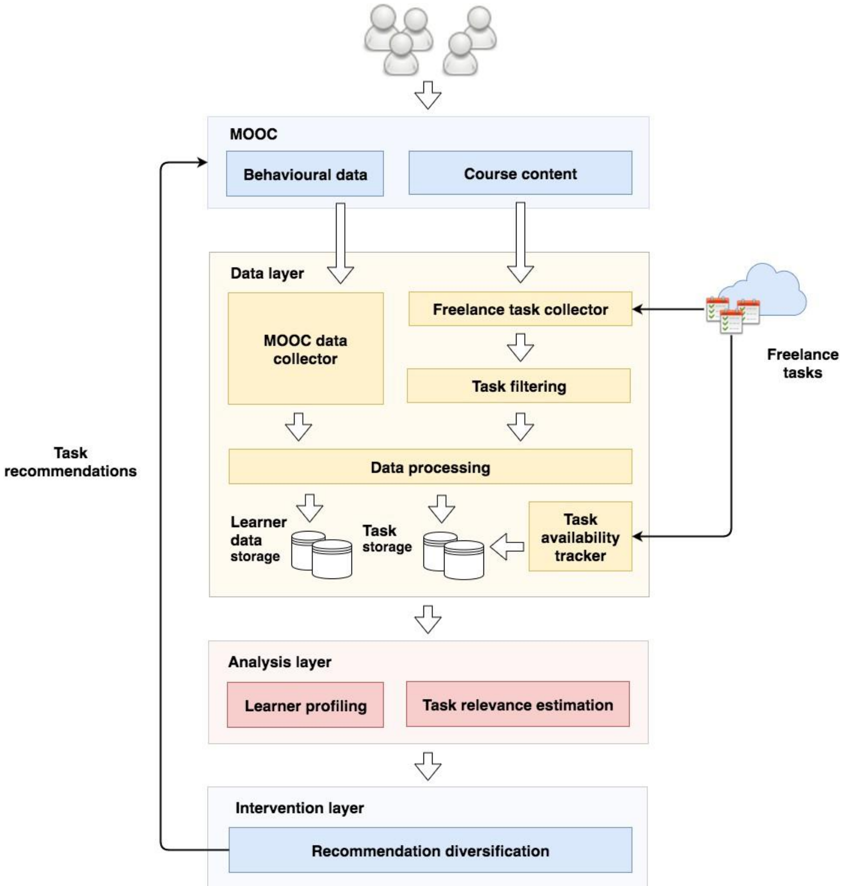


Figure 1: Rec-$ys architecture.

Implications: Can MOOC learners be paid to learn? We set out to provide a first answer to this question in the context of a data analysis MOOC. We found that indeed, work tasks can be solved accurately and in high quality by a considerable percentage of learners that attempt it. Based on the work presented here, we will explore several promising directions: (i) experimental setups that allow us to further investigate the causal relationship between real-world tasks and learner engagement, (ii) the suitability of more complex tasks for MOOC learners, (iii) the acceptance of the "learners can be earners" paradigm in different populations, and (iv) setups that aid MOOC learners to take the first steps in the paid freelance task world

Acknowledgments: This abstract is based on a journal article [1] that was accepted by IEEE Transactions on Learning Technologies and a poster paper [2] that was accepted by the 2017 Conference on Learning Analytics and Knowledge (LAK 2017).

Guanliang Chen's work is supported by the Extension School of the Delft University of Technology.Dan Davis' work is supported by the Leiden-Delft-Erasmus Centre for Education and Learning.

Resources:

1. Guanliang Chen, Dan Davis, Markus Krause, Efthimia Aivaloglou, Claudia Hauff and Geert-Jan Houben (2016). Can Learners be Earners? Investigating a Design to Enable MOOC Learners to Apply their Skills and Earn Money in an Online Market Place. IEEE Transactions on Learning Technologies.
2. Guanliang Chen, Dan Davis, Markus Krause, Claudia Hauff and Geert-Jan Houben (2017). Buying Time: Enabling Learners to become Earners with a Real-World Paid Task Recommender System. In Proceedings of the 7th International Conference on Learning Analytics and Knowledge, pp. 578--579, Vancouver, Canada. LAK'17.

Long-term effects of study-choice meetings, online personal goal-setting, and an academic stretch goal on student performance

Michaéla Schippers, Ad Scheepers, Niek Hoogervorst
Rotterdam School of Management, Erasmus University, Rotterdam, Netherlands

Purpose: Numerous studies have focused on determining factors of academic performance. Different categories of determining factors can be distinguished (e.g., personality traits, motivational factors, contextual factors). Psychosocial contextual factors appear to be the strongest correlates with academic performance (e.g., see Richardson, Abrahams, & Bond, 2012). However, most academic psychosocial contextual interventions are administered just before or at the start of an undergraduate program and usually effects are measured in or after the first academic year.

Although several of these studies show significant effects on the short-term, it is unclear if the interventions have lasting effects. We will present the long-term academic-performance effects of three major psychosocial contextual interventions, administered just before or in the first months of a 3-year bachelor program.

Method: We investigated three early interventions (study-choice meetings, online personal goal- setting, and an academic stretch goal) administered to six different first-year undergraduate cohorts (N = 4576) at a large European business school over six years. Long-term academic effects were considered.

Results: Results show that after three and four years participants still benefited academically to a significant degree from the interventions at the start of their degree program. Interestingly, different

personality characteristics appeared to moderate performance at different stages of the study program.

Conclusion: Early interventions with undergraduate students can lead to stable and robust academic effects over the long term. Moderator effects on short term academic performance can be different than moderator effects on long term academic performance.

"WHY DO I NEED THIS?" - HELPING STUDENTS UNDERSTAND MARKET DEMAND FOR INDIVIDUAL SKILLS

Scott Harrison, Christian Weber, Sisay Adugna Chala, Madjid Fathi
University of Siegen, Institute of Knowledge Based Systems & Knowledge Management (Germany)

Purpose: Students in higher education are continuously facing pressure to understand many dimensions of the curriculum the relevance of topics or lessons and how they relate to their personal goals and the wider world. This raises questions about the quality, time and availability of their education, fitting their interests and their market motivated needs to the right skills, in the right context, at the right time, to attract the "right" employer.

Previous research has indicated that for most students "who could benefit from such occupations are unaware of job openings, the salaries they offer, or the credentials needed to secure them" (Aspen Institute 2013). Selecting and engaging in the course options in a given curriculum, students begs the question "Why do I need this?". On the other hand, teachers are challenged to develop a curriculum that is both current and relevant to the needs of students. From the labour market information perspective, it has its own challenges when trying to integrate into the learning environment. For example "measuring the rate of technological change as it affects the labour market has proven difficult, and labour market policy needs to be based on more than the casual empiricism behind the claim that the world is changing faster than ever. Researchers have devoted considerable effort to address these issues, much of it summarized in a companion document to this report" (Handel 2012).

To address this duality, the matching process of education data to job market data needs to be considered. A study by the OECD (2012) on the trends in job skill demands in OECD countries highlights that "the long-term trend has been toward jobs requiring more education and cognitive skills, but the rate and timing of changes, the precise level and kinds of skills in demand, and the drivers of change are matters of debate and are often poorly understood". While this is still true for recent research, it is even less understood by teachers and students within higher education institutions. Another key study by CEDEFOP outlines the importance of this information, and has compiled a comprehensive case study report, outlining Labor Market Information systems (LMI) used in 11 European countries. It concluded that “LMI should be well-integrated in a career learning process that promotes the development of reflexive career identities and autonomous exploration of career information. LMI should not be seen as a stand-alone tool, as is frequently the case” (CEDEFOP 2016 pg. 81). As such, matching today’s higher education and job market requirements can provide important insights and promote a better job matching process. Learning analytics thrives to improve the learning environment for the student, answering what, when and how well do I learn? Job market analysis approaches from another direction; what skills are needed to fill which job role. The match finally opens the question, what learning outcome can satisfy which skill?

Design: This paper proposes a different path to address the matching, using a new approach to utilize information to select and improve courses. Students scan and evaluate the course description, outcomes and requirements to select the best set of courses in a curriculum to match their vision of their future job. Organisations scan and evaluate job, job roles and skill-set descriptions to match their need to a future employee, informed by sources such as ESCO, ISCO and WageIndicator. While job descriptions are well explored, the student vision of their future job is less known, especially in the presence of labor market data.

We propose a novel student-focused approach to matching and improving both course learning outcomes in line with the market requirements and the student expectation. Teachers get the opportunity to select, in line with the European skill and competency standard, which job roles, competencies and skills are covered by their courses, addressing the learning outcomes. Students will gain context to the skills being taught in the curriculum when provided job data, enabling them to rate courses and their skill match using a simple traffic light rating. A green indicator is given by the student when they can see the value in the skills being taught are in line with job market demands; yellow lights indicate students can see value in the skill, but it's not inline with market demand, and red indicates the students perceive no reason for learning the skill. Following up, students can then rate the match, using traffic lights in different stages of the course, uncovering first a market expectation mismatch, then a course expectation mismatch, and finally a course quality (or learning skill) mismatch. The new approach will be introduced and explored from a learning analytics perspective, to help to give a new answer to “Why do I need this?”

Results: The study is in the formative stages of development, however, the key results that are available focus more on the technological design and integration of mixed data sources to enable the implementation of such a system.

Implications: The expected implications of this research are 1) Better contextualisation of lessons being taught - it is a reasonable expectation that through better context driven learning environments, students will gain a better understanding how specific skills being taught have real world implications. 2) Increased student engagement in skills development - A follow on effect from contextualisation is the engagement process itself. 3) Improved conceptualisation of the student's job focus - Skills and job roles can fit to different jobs and enable to analyse the feedback towards isolating and recommending a job perspective of students. In all, we foresee this approach facilitating better provision of labor market information to students, while also developing data sets that enable us to understand what effect the information has on future outcomes for the students.

Acknowledgments: The authors would like to acknowledge the financial support of the Eduworks Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Program.

Resources:

1. CEDEFOP (2016). Labour market information and guidance. Luxembourg: Publications Office. Cedefop research paper; No 55. http://dx.doi.org/10.2801/72440
2. Aspen Institute (2013) A Guide for Using Labor Market Data to Improve Student Success. https://eric.ed.gov/?id=ED541431
3. Handel, M. (2012), "Trends in Job Skill Demands in OECD Countries", OECD Social, Employment and Migration Working Papers, No. 143, OECD Publishing, Paris. http://dx.doi.org/10.1787/5k8zk8pcq6td-en

Methods and Data 2 - Self-regulated and adaptive learning

Analysing adaptive learning platforms utilizing domain ontologies: Searching for analytical implications.

Réka Vas, Christian Weber

Corvinus University of Budapest, Hungary

University of Siegen, Germany

Purpose: Competition in e-learning solutions is increasing at an alarming rate, in line with the frequent and diverse changes in the expectations of both learners and the labour market. Moreover, different students have different learning styles. Some may follow a linear reasoning process when solving a problem, others may prefer making intuitive leaps. Some may deal well with theories, others may learn through examples or experiments. At the same time, most curricula found even in online training programs or learning systems are designed in a linear fashion, where content is structured in such a way that it apparently makes sense if process the curricula chapter by chapter. Naturally, this does not suit the needs of every single learner. This work aims at introducing the STUDIO learning and knowledge assessment method and summarizing major concerns on how data on user behaviour could be or should be analysed to provide better learning experience.

Design: To enable personalized learning, the STUDIO learning system – making use of domain ontologies and an adaptive test engine – has been designed in such a way that it can compile and re-compile personalized self-assessment lessons (including both learning materials and assessment questions) following the guidance of the domain ontology – this way breaking with traditional linear curricula structuring. In the first stage, the learner's knowledge is tested and evaluated in STUDIO to identify those concepts (of the domain of interest) where the test candidate has deficiencies. Based on the result of this first self-assessment test, a set of personalized learning materials is provided with guidelines on how the learner should "walk through" the ontology structure in the process of learning. Access is provided to the learning material of those concepts where the student incorrectly answered the related test question. Final results are visualized on a graph representation of the domain ontology, providing a tailored view of the interconnected concepts. This graph applies ample colouring to label the current performance of the learner to enable the exploration of the learner's knowledge gaps in the context of the domain. Since results are represented using the ontology visualization tool of STUDIO, not only concepts but also their interdependencies are presented to define the proposed paths of learning.

To follow and analyse the learner's interaction with STUDIO, three groups of data sets are collected:

1. Performance data: includes data on test results, combining different indicators to track the overall rate of passed concepts in comparison to the total number of concepts or to the number of tested concepts.
2. Assessment data: includes data concerning which question related to a given concept was asked, when this question was asked and what the answer was from the four possible answers.
3. Interaction sequence data: includes data concerning with which elements of the interface the user interacted and when. In sequence and joined with the other data sets, it traces self-assessment, result visualization and interaction

with the learning material. Selected interactions, as viewing learning materials, are aggregated to counter.

Results: The STUDIO system (method) was applied in a variety of domains, ranging from accompanying organisational processes for non-formal learning to supporting formal education for a blended learning approach in higher education classes (HRIS, MIS). Following the classification of Hoppe (2017) a process-oriented interaction analysis was conducted, fusing logged interaction data for analysis. Throughout the use of the system, data was collected across different data sources, tracking the test progress and results but also in which sequence and how long a learner interacts with which component of the system.

Performance focused correlation analyses for the blended learning scheme, have shown only minor indications of performance gains in the middle and final examinations of classes (MIS) in comparison to years without the system. Learned clusters of low, middle, and highly performing students can only weakly explain changes found in their performance.

Revisiting the results, it is yet unclear which factors may fuel better learning outcomes or e.g. what role the motivation of the learner plays in reaching continuous performance improvement (even in those cases where the number of learning material views is very low). Additional analysis is needed to uncover the potential impacts of the ontology driven curriculum structure. Deducting what further analyses to use and how to modify the experiments for an improved data collection would benefit from the expert vision and lessons learned of the learning analytics community.

Implications: The STUDIO system describes a unique approach to structure and guide learning following a domain ontology guiding the learner through an adaptive long-term learning process. Exploiting the domain ontology and tracking the user interaction and performance over time, the system offers a good potential to combine observations of learning and the specific composition and representation of the domain to derive new model driven insights into learning processes. Yet, as of the nature of the data and the circular character of the STUDIO method, the question prevails how the different data sources are analysed best and how experiments should be designed to account for the circular system character.

One consideration regarding experiment design is how to filter out outliers or unmotivated learners. Analysis of both individual learning curves or system level learning curves is biased if learners who are not motivated to reach real improvement, or learners who use external sources to learn are not filtered out.

Another major consideration regarding experiment design is how to measure the effect of the curriculum structuring approach of STUDIO. It is expected that the explicit interconnection of concepts in the domain of interest and the overview, provided by ontology visualization, should provide better understanding of the curriculum. At the same time, it is not clear when these tools really trigger improvement in learning results. It might happen that benefits only appear on the long term (e.g. an improved long-term retention rate) and not in the course of interacting with the system.

Acknowledgments: The authors acknowledge the financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission’s 7th Framework Program.

Resources:

1. STUDIO system demo: https://www.youtube.com/watch?v=a3RZcV1rOVk

Investigating the relationships among self-regulation, approach to learning, goal orientation, LMS activity and academic performance.

Vladimer Kobayashi[1], Prathyusha Sanagavarapu[2], Catherine Zhao[3], Stefan Mol[1], and Gábor Kismihók[1]

[1]*University of Amsterdam,* [2]*Western Sydney University, and* [3]*University of Sydney*

Purpose: We investigated the simultaneous effect of self-regulation (Barnard, Lan, To, Paton, & Lai, 2009), approach to learning (Biggs, Kember, & Leung, 2001), goal orientation (Elliot & Murayama, 2008) and Learning Management System (LMS) (e.g. Blackboard) activities of students on their final grades (as a measure of academic achievement). We postulate that academic achievement can be better explained if we take several factors at the same time. Apart from incorporating observable actions carried out by students (e.g. goals they set, actions they take to achieve the goals, and their LMS activities), intrinsic and behavioural characteristics of students should also be taken into consideration since these characteristics ultimately influence their actions. Also, in this study we aim to find out which of the intrinsic and behavioural characteristics explain students' LMS activities.

Design: The subjects were college students. For data collection, questionnaires were used to measure self-regulation, approach to learning, and goal orientation. Moreover, students were encouraged to use a goal setting application that enabled them to set, manage, and monitor individual goals. Through the goal setting app, we collected actual goals set by students. LMS activity data were gathered at the end of the course.

Results: We found the following relationships between constructs and LMS activity:

For goal orientation: (1) Mastery avoidance positively correlates with the number of characters posted in discussion forum; (2) Mastery approach is positively correlated to LMS site access halfway through the course; (3) Performance avoidance is positive correlated to LMS site access halfway through the course and negatively correlated to LMS site access toward the end of course. For approach to learning: (1) Deep strategy is positively correlated to both content and item access in LMS.

We subsequently built a model based on decision tree technique that predicts academic performance. Both LMS and questionnaires data were used to build the model. Task strategy (from self-regulation), performance approach goal orientation, content item access, and site access halfway through the course are good predictors of academic performance. Using Bayesian network, it was further revealed that mastery avoidance and task strategy directly influences the final mark, that is, mastery avoidance adversely affects academic performance whereas task strategy positively affects it. There is an indication that students who set goals are more likely to get higher grades than students who do not set goals.

Implications: Our findings showed that some LMS derived measures (e.g. frequency and timing of LMS access, participation in discussion forums, and item content access) can be used as indicators of academic success (Arnold & Pistilli, 2012). Task strategy and goal orientation of students also influence their academic achievement. Moreover, observed LMS activity are partly explained by students' approach to learning and goal orientation. We conclude that to better predict academic achievement researchers should consider intrinsic and behavioural characteristics of students as well as the specific actions they take (such as goals they set and their LMS activities). This also has ramifications in the way we detect students at risk of failing and the type of intervention we employ to help students. LMS data may provide a quick way to diagnose students who are about to fail and behavioural characteristics enable us to determine appropriate interventions.

Acknowledgments: We are grateful for the financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Programme.

References:

1. Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (pp. 267–270). New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330666
2. Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self- regulation in online and blended learning environments. *The Internet and Higher Education*, *12*(1), 1–6. https://doi.org/10.1016/j.iheduc.2008.10.005
3. Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, *71*(1), 133– 149. https://doi.org/10.1348/000709901158433
4. Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100*(3), 613.

The effect of adaptive practicing on the relation between students' summative grades and following learning activity

N. van Halem, C. van Klaveren, & I. Cornelisz
Amsterdam Centre for Learning Analytics (ACLA),
Faculty of Behavioural and Movement Science, Vrije Universiteit Amsterdam

Purpose: While adaptive practicing is increasingly common in education, no actual evidence on the effects on the learning process is established yet. This study investigates the effect of adaptive versus static practicing on students' learning activity, and its relation with obtained summative grades.

Design: This study takes place over the course of one school year, in the context of the lower grades of secondary education and the courses biology, economics, and history. Students were randomly assigned to adaptive versus static practicing within an existing digital learning environment. Analyses were disaggregated into quartiles of student's average achievement level and based on a longitudinal hierarchical regression model (N = 606), yielding the proportion of variance between and within students (over time).

Results and implications: Results suggest that adaptive practicing positively affects learning activity amongst certain – mostly high-achieving - students. Future research is required to fully explain these results and optimise adaptive practicing in education.

Follow the Successful Crowd: Raising MOOC Completion Rates through Social Comparison at Scale

Dan Davis[1], Ioana Jivet[2], René Kizilcec[3], Guanliang Chen[1], Claudia Hauff[1], Geert-Jan Houben[1]

[1]TU Delft Lambda Lab, [2]Open University Netherlands, [3]Stanford University

Purpose: Social comparison theory asserts that we establish our social and personal worth by comparing ourselves to others. In in-person learning environments, social comparison offers students critical feedback on how to behave and be successful. By contrast, online learning environments afford fewer social cues to facilitate social comparison. Can increased availability of such cues promote effective self-regulatory behavior and achievement in Massive Open Online Courses (MOOCs)? We developed a personalized Feedback System that facilitates social comparison with previously successful learners based on an interactive visualization of multiple behavioral indicators. Across four randomized controlled trials in MOOCs including more than 30,000 learners, we find: (1) the availability of social comparison cues significantly increases completion rates, (2) this type of feedback only benefits highly educated learners, and (3) learners' cultural context plays a significant role in their course engagement and achievement.

Design: A mechanism for increasing access to higher education content, MOOCs have afforded millions of people worldwide the opportunity to learn for little or no cost. To achieve this unprecedented scale, MOOCs provide the same material to all learners, no matter what background, motivation, and skill set they possess. Yet this approach falls short of leveraging the technical possibilities of contemporary educational resources to offer learners personalized support, such as giving guidance to learners who are unskilled at regulating their learning process over several weeks to achieve mastery. Low course completion rates (typically between 5-10%) highlight the need for additional instructional support in MOOCs.

While many learners have no intention to complete MOOCs and instead use them to fulfill alternative needs, the majority of learners who are motivated and committed to complete the course still fail to achieve their goal [1, 2]. Most learners report that they could not find the time to keep up with the course, a challenge that is related to insufficient self-regulatory abilities [3, 4]. Self-regulated learning (SRL; i.e., the ability to plan, monitor, and actively control one's learning process) is associated with a higher likelihood of achieving personal course goals in MOOCs, including course completion [5, 6]. However, the current design of MOOCs does not support learners to engage in SRL [7]. In particular, most MOOC platforms do not provide learners with personalized feedback beyond grades [8], and thus, learners may not know if their engagement in the course is conducive to achieving their learning goals.

We introduce a system that facilitates social comparison to help learners regulate their learning behavior to support course completion. According to social comparison theory [9], people establish their social and personal worth by comparing themselves to others. In addition to evaluating the impact of providing learners with personalized feedback, we further examine the potential of adjusting the framing of the feedback to match learners' cultural context. Framing feedback in a way that is consistent with the norms and achievement-based motivation of learners' cultural context is expected to support internalization and behavior change. Prior work has observed differences in the way learners from different countries and cultures interact with MOOCs [10, 4, 11]. We define cultural context based on two established country-level cultural dimensions: individualism by Hofstede et al. [12] and tightness by Gelfand et al. [13]. We explore the extent to which insights from the social comparison and cultural psychology literature can be translated to support learners in MOOCs. We evaluate how to offer feedback based on social comparison in an online learning environment. To this end, we design, develop, and empirically evaluate a personalized and scalable feedback system that presents MOOC learners with a visual comparison of their behavior to that of their "successful" peers, that is, those who completed the course in the past.

Results: Our work extends prior research by testing a theory-informed technological solution in a large and diverse population (i.e., MOOC learners) for a prolonged period of time. These are our main findings:

- Personalized social-comparison feedback increases course completion rates.
- Only highly educated learners benefit from this kind of feedback.
- Course engagement and achievement varies by cultural context: learners in countries with a "loose" culture outperform those in countries with a "tight" culture.

Implications: We find that the Feedback System primarily benefits highly educated learners, although the system was envisioned to support those who struggle with self-regulation. This suggests a new challenge for MOOC researchers and designers to make targeted interventions that support learners who need more support. As online courses can be culturally diverse learning environments, we investigated how the Feedback System could be adapted to resonate with learners from different backgrounds. Aside from our intervention, we found that learners from loose cultures consistently outperformed learner's tight cultures in terms of course engagement and final grades. In light of the two sources of heterogeneity we identified, future MOOC interventions may be strengthened by personalization based on learners' prior education level and cultural context.

Acknowledgments: This abstract is based on a full research paper [14] that was presented at the 2017 Conference on Learning Analytics and Knowledge (LAK 2017). Dan Davis' work is supported by the Leiden-Delft-Erasmus Centre for Education and Learning. Guanliang Chen's work is supported by the Extension School of the Delft University of Technology.

Resources:

1. R. F. Kizilcec, C. Piech, and E. Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In LAK '13, pages 170–179. ACM, 2013.
2. R. F. Kizilcec and E. Schneider. Motivation As a Lens to Understand Online Learners: Toward Data-Driven Design with the OLEI Scale. TOCHI, 22(2):6:1–6:24, 2015.
3. S. Zheng, M. B. Rosson, P. C. Shih, and J. M. Carroll. Understanding student motivation, behaviors and

perceptions in moocs. In CSCW ’15, pages 1882–1895, 2015.

4. R. F. Kizilcec and S. Halawa. Attrition and achievement gaps in online learning. In L@S ’15, pages 57–66, 2015.
5. R. F. Kizilcec, M. P´erez-Sanagust´ın, and J. J. Maldonado. Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. Computers & Education, pages 18–33, 2016.
6. Littlejohn, N. Hood, C. Milligan, and P. Mustain. Learning in moocs: Motivations and self-regulated learning in moocs. The Internet and Higher Education, 29:40–48, 2016.
7. Margaryan, M. Bianco, and A. Littlejohn. Instructional quality of MOOCs. Computers & Education, 80:77–83, 2015.
8. D. Davis, G. Chen, I. Jivet, C. Hauff, and G. J. Houben. Encouraging Metacognition & Self-Regulation in MOOCs through Increased Learner Feedback. In Workshop on Learning Analytics for Learners, 2016.
9. L. Festinger. A theory of social comparison processes. Human relations, 7(2):117– 140, 1954.
10. P. J. Guo and K. Reinecke. Demographic differences in how students navigate through MOOCs. In L@S ’14, pages 21–30, 2014.
11. Z. Liu, R. Brown, C. F. Lynch, T. Barnes, R. Baker, Y. Bergner, and D. McNamara. Mooc learner behaviors by country and culture; an exploratory analysis. In EDM ’16, pages 127–134, 2016.
12. G. Hofstede, G. J. Hofstede, and M. Minkov. Cultures and organizations: Software of the mind. 1991.
13. M. J. Gelfand, J. L. Raver, L. Nishii, L. M. Leslie, J. Lun, B. C. Lim, L. Duan, A. Almaliach, S. Ang, J. Arnadottir, et al. Differences between tight and loose cultures: A 33-nation study. Science, 332(6033):1100–1104, 2011.

14. Davis, D., Jivet, I., Kizilcec, R. F., Chen, G., Hauff, C., & Houben, G. J. (2017, March). Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 454-463). ACM.

Gamification in interactive learning environments

David Stap, Bert Bredeweg and Natasa Brouwer
University of Amsterdam, Informatics Institute & Education Service Centre

Purpose: Gamification uses game-design elements and game principles in a non-game context [3]. Typical gamification methods let users earn Experience Points (XP) and unlock badges. The most important goal of gamification is increasing motivation [5]. Intrinsic motivation is a particularly strong type of motivation [8]. Attempts have been made to use social and other motivations as intrinsic motivation in a gamification context [6, 7]. Research showed that badge-based achievement systems can have a highly significant positive effect on the quantity of students' contributions [2]. It was also shown that students who completed a gamified experience obtained better scores in practical assignments and in overall score [4]. However, it remains unclear in what context gamification techniques are best employed: studies find mixed results when employing similar game mechanics [9].

DynaLearn is a browser-based interactive learning environment (www.DynaLearn.nl), where users learn by qualitative modelling [1]. Learners develop their ideas by creating and simulating such models. DynaLearn is typically used in secondary education for courses such as biology and physics. Evaluation studies have shown that this learning environment improves learners' system thinking.

The aim of research presented here is to add experience value to working with learning by modelling environments (such as DynaLearn) by providing Learning Analytics based feedback to students using gamification techniques, with the main goal of increasing student motivation.

Design: Three game-mechanics have been implemented: badges, leaderboards and lives [7]. These mechanics are integrated into DynaLearn: learners can inspect the mechanics via the user bar at the top of the screen.

Badges are awarded to users for certain types of behaviour or achievement, their goal is to provide feedback and motivation to students. A total of 16 badges were developed and implemented. The selection of badges and their application in the online module was optimized to ensure a constant stream of feedback while doing the assignment. When a learner receives a new badge, the badges button blinks to notify the learner.

A leaderboard displays ranked user scores. The underlying idea is that this drives competition, which proved to be stimulating for certain types of learners. Furthermore, the leaderboard provides users with an overview of their relative performance. Every time a learner performs an action, the score gets updated on the leaderboard in DynaLearn. When the position on the leaderboard changes, a blinking leaderboard icon notifies the learner. A maximum score can be obtained by minimizing mistakes and minimizing total number of correct actions, i.e. a model builds according to the specifications of the assignment results in a high score.

The goal of the life mechanic is to let learners deliberately think about their next action, as opposed to adding components without much thought. For every learner performed action, the life mechanic checks whether this action is consistent with a norm model. If the action is inconsistent (e.g. an incorrect name), the learner loses a life. All learners start with 3 lives. The life mechanic should trigger learners to carefully examine next actions, but allow for some mistakes. If a learner loses all lives, the 'game is over' and the learner must restart the modelling effort.

Evaluation study – Design

For evaluation, a physics assignment was created and administered in a real classroom (Havo3, 9^{th} grade). The goal of the assignment was to model in what way the gravitational forces between the earth and sun are impacted by the changing mass of the sun. To successfully make the assignments, learners understand the causal relations concerning gravitation. The solution to the assignment consisted of 15 elements, including entities, configurations, quantities, proportionalities (positive and negative), and quantity values [1].

Two groups participated in the experiment: the control group (n = 11), who used the unaltered DynaLearn environment, and the treatment group (n = 24), who used DynaLearn with gamification elements. Both groups of learners were already familiar with the software. assignment. User modelling behaviour was tracked using Learning Analytics techniques to find differences in behaviour between the groups. A survey was conducted with the goal of measuring motivation, attitude about gaming, and situational awareness. Additionally, the learner created models were compared to see if there were differences in performance. Finally, some informal observations about the two sessions were noted.

Results: Figure 1 shows the type and number of actions (create, modify, and delete) performed by control and treatment group. As can be seen, the treatment group performed less delete actions (this difference was significant). The treatment group also had a higher value regarding the standard deviation on total actions, indicating an uneven activitydistribution.

Figure 1. Number of actions by ID (left: control group, right: treatment group)

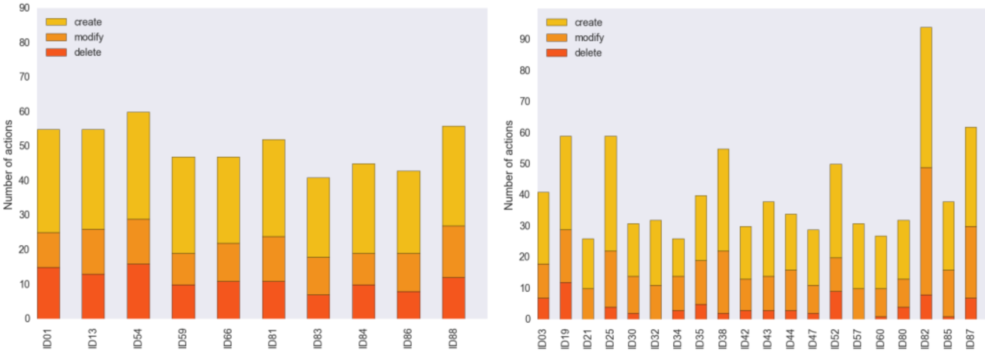
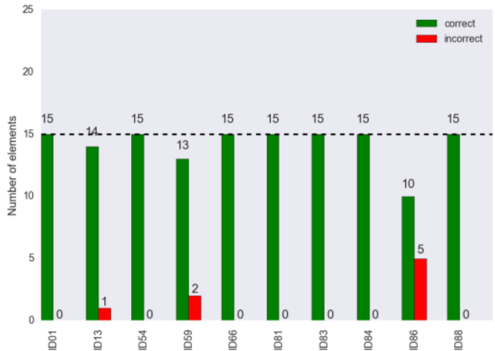
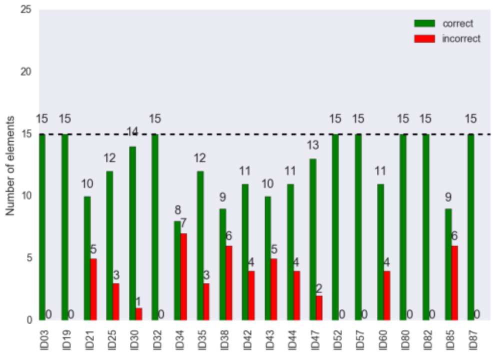


Figure 2 shows the distribution of correct and incorrect components in the models created by the learners. The dotted line in the figure indicate the maximum number of correct elements. The learners in the treatment group perform worse than learners in the control group: they make more mistakes.

Figure 2. Correctness of final model (left: control group, right: treatment group)





The survey revealed that learners from the treatment group scored higher on Q3 ('I would like to have another such lesson') compared to the control group.

Informal observation showed that a significant portion of learners in the treatment condition seemed very occupied with the leaderboard. Some checked their score and ranking after every move, others shielded their laptop screen to prevent others from copying their models. A few learners got somewhat annoyed with the life mechanic, because they had to restart on a blank canvas several times. This may have resulted in a lower level of motivation. Finally, most learners did not seem very interested in their badges.

Implications: The treatment group performed a significantly lower number of delete actions compared to the control group, while learners that perform a high number of delete actions are more likely to create superior models when compared to learners that perform a lower number of delete actions. Moreover, the analysis showed that the treatment group scored higher on Q3, indicating that they enjoyed the game mechanics. Together these results seem to suggest that gamification resulted in more fun but inferior learner created models (and possibly, inferior learning).

Future research should investigate the reasons why gamification resulted in inferior learner created models. An explanation could be that gamification in this social scenario, although engaging learners, resulted in too many distractions. As a result, the cognitive effort of learners is lost to the game mechanics.

Resources:

1. B. Bredeweg, J. Liem, W. Beek, F. Linnebank, J. Gracia, E. Lozano, M. Wißner, R. Bühling, P. Salles, R. Noble, A. Zitek, P. Borisova, and D. Mioduser. Dynalearn – An intelligent learning environment for learning conceptual knowledge. AI Magazine, 34(4):46-65, 2013.
2. P. Denny. The effect of virtual achievements on student engagement. In Proceedings of the SIGCHI conference on human factors in computing systems, pages 763-772. ACM, 2013.
3. S. Deterding, D. Dixon, R. Khaled, and L. Nacke. From game design elements to gamefulness: defining gamification. In Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments, pages 9-15. ACM, 2011.
4. Dominguez, J. Saenz-De-Navarrete, L. De-Marcos, L. Fernandez- Sanz, C. Pages, and J. Martinez-Herraiz. Gamifying learning experiences: Practical implications and outcomes. Computers & Education, 63:380-392, 2013.
5. Glover. Play as you learn: gamification as a technique for motivating learners. 2013.
6. Hamari and J. Koivisto. Social motivations to use gamification: An empirical study of gamifying exercise. In ECIS, pages 105-117, 2013.
7. Hamari, J. Koivisto, and H. Sarsa. Does gamification work? – a literature review of empirical studies on gamification. In System Sciences (HICSS), 2014 47th Hawaii International Conference on, pages 3025-3034. IEEE, 2014.

8. R.M. Ryan and E.L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American psychologist, 55(1):68, 2000.
9. Seaborn and D.I. Fels. Gamification in theory and action: A survey. International Journal of Human-Computer Studies, 74:14-31, 2015.

The Higher Education Enrollment Decision: Bayesian Learners versus Bad News Ignorers

Chris van Klaveren, Karen Kooiman, Ilja Cornelisz & Martijn Meeter
Vrije Universiteit Amsterdam

Purpose: The empirical literature shows that students suffer from the self-serving bias (SSB), which means that expectations of secondary/high school students about their own future study success are overly positive, either because they are overconfident or because they are insufficiently informed (Zafar, 2011). This may result in a study choice that does not fit well the student's cognitive capabilities, and eventually may lead to dropout in the first year, because once being enrolled students will experience that the required skills necessary for academic success are lacking.

To overcome the potential negative effects of the *SSB,* Dutch secondary school students who applied for study in higher education, but are not yet enrolled in the educational program, receive predictions on their future study success. These predictions are obtained by estimating the probability of study success for the previous student cohort conditionally on study type, proficiency test scores on study-related content and the moment of application. Based on these estimation results out-of-sample predictions are performed for students who applied for a study but still have to make the higher education enrollment decision. These out- of-sample predictions are communicated to students textually and visually, such that students can update their beliefs about their future academic success and can make a well-informed higher education enrollment decision, which in turn is intended to improve the quality of their study choice and result in lower dropout probability in the first year.

Design: A randomized field experiment was conducted among secondary school students who applied in 2016 for a law or social sciences study at the Vrije Universiteit Amsterdam, to evaluate the effects of providing students with information on their future study success. The experiment begins when secondary school students attend an information day, which is organized by Dutch university institutes, that has the objective to improve the quality of the match between the higher education enrollment decision of students and the offered content in Dutch higher education programs. Students who gave their informed consent (313 of the 627 applicants) were randomly assigned to a control group (who received no information) or an intervention group (who received information). This study evaluates if student expectations before enrollment are, on average, overly positive and if providing students with information on study success improves their higher education enrollment decision.

Results: The empirical findings confirm that students tend to be overly positive about their own performance and the prevalence of the self-serving bias is more pronounced for students for whom we predict lower study success. The higher education enrollment rate of students who received feedback was significantly higher with about 25 percent. The self-serving bias does not drive the higher education decision, but instead the results indicate that the enrollment rate tends to be lower if the returned probability of study success was lower than 57 percent, and tends to be higher if if the returned probability of study success was higher or equal than 57 percent. An explanation for these results could be that Dutch students are generally graded on a scale from 1-10 and a grade higher than 5.5 may generally be perceived as 'positive news' as it resembles a sufficient mark. Even though returning information to students on study success affects enrollment rates positively, it does not significantly affect student performance or enrollment status once being enrolled in the educational program. Finally, a follow-up questionnaire shows that students do not ignore bad news, but update expectations in a direction that is consistent with the behavior of Bayesian updaters.

Implications: This study supports existing empirical evidence and shows that the higher education decision can be influenced by a relatively low cost intervention. At the same time, the results also confirm that long-term outcomes are not significantly improved. Several explanations can be given for this. First of all, programs organized by higher education institutes tend to focus on a limited set of higher education programs (generally one educational program). As a result, treated students receive additional information about one study, but not about alternative studies that also belong to the relevant choice set. This also holds for this study: students receive predictions on study success for one study, but not for other relevant studies in the choice set, that makes it difficult for students to act optimally based upon the information received. Secondly, most interventions are low stake interventions, which may prevent positive and substantial behavioral effects on the long-term.

Finally, we would like to point out, at least for the Netherlands, that first-year switching behavior may represent the iterative process of making an optimal higher education decision. Secondary schools and higher education institutes organize events which should enable students to make a well-informed choice about which field of study matches their personal preferences and cognitive capabilities best. Moreover, students who enroll in higher education should theoretically be cognitively capable of doing so, because they have successfully finished pre-university secondary education. Therefore, it is questionably if higher education institutes should invest substantial amounts of money to prevent first-year dropout, as preventing first-year switching behavior could also potentially result in a suboptimal higher education decision.

Session 1 - Issues of privacy and ethics

Workshop, panelists: Niall Sclater, Fay Kartner, Nynke de Boer

Session 2 Early warning systems and predicting student success or failure

Sense Making of Student Analytics Development and Application of an Early Warning Model to prevent Bachelor Dropout

Theo C. Bakker
Vrije Universiteit Amsterdam

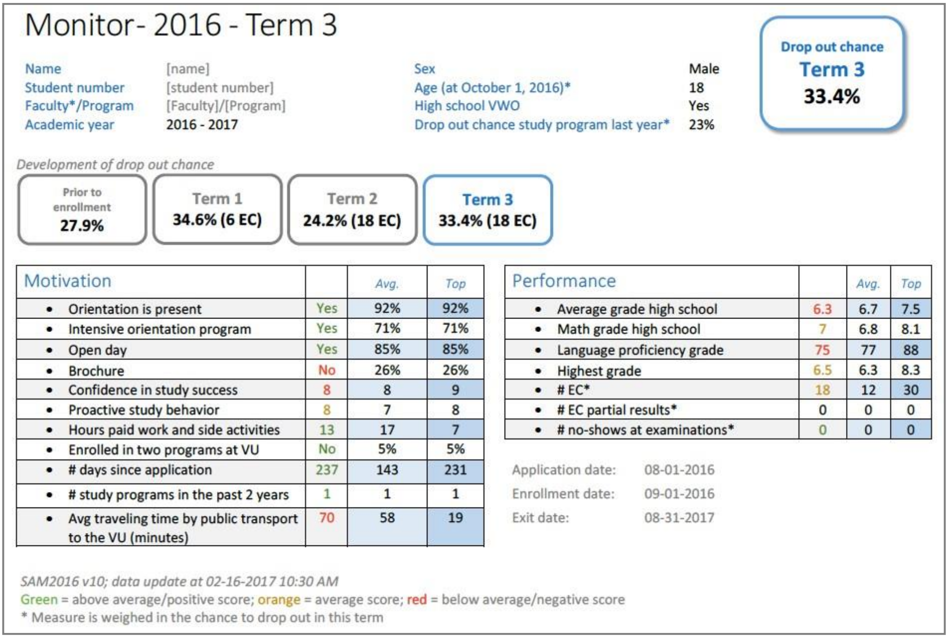
Purpose: This presentation has two goals: 1) To share a three-year process of policy making, stake holder management, student data collection, engineering, analysis, and modelling. This resulted in a university wide project and dataset containing data from 25 sources and 65.000 students and a model for student retention. 2) To present the application of the model within an early warning system used within an experiment by student counsellors and tutors for counselling of first-year bachelor students of a Dutch university.

Design: The data for the model were originally collected as part of a student analytics project which started in 2014. The main objective of the data collection was to improve enrollment, progression and graduation of students. This university wide project derived most of the data from transactional/ administrative systems, both internally from the university, as externally from national data sets. Exceptions were a matching and student satisfaction questionnaire. Furthermore, data was enriched with public data on demographic background, open data on secondary schools and traveling times by public transport from the secondary school to the campus.

The main data set, hereinafter ’the data set’, consisted of student data spanning 7 academic years of data (enrollments in 2010-2016) of 28.000 first year bachelor students of a Dutch university in 54 study programs of 11 faculties. Each record in the data set defined one enrollment in a specific academic year and study program.

Based on student data of about 5.000 bachelor students of the academic year 2014 and 2015 a prediction model (GAM) was developed to predict student retention after the first year. The model changes during the first year as it uses available predictors to adjust the chance of dropping out: before the study and after each of the 6 educational terms of the first year. Predictors were derived from data collected before the study (such as demographics, previous education, orientation, matching, application data) and during the study (language proficiency test and study results).

An experiment was designed to study the effects on first-year student retention of college counseling based upon an early warning system. The measure for student retention was the dropout rate of first-year students after one year. In September 2016 students of three faculties were asked for their permission to use their data, apply the model and use the outcomes in student counselling, following a procedure developed with the privacy office of the university. Participants were randomly assigned to a test and control group. Student counsellors and tutors were provided on a weekly basis with a ‘student monitor’ for each participating student to be used in student counselling (*see below*). About 800 students granted their permission.



Monitor- 2016 - Term 3

Name	[name]	Sex	Male
Student number	[student number]	Age (at October 1, 2016)*	18
Faculty*/Program	[Faculty]/[Program]	High school VWO	Yes
Academic year	2016 - 2017	Drop out chance study program last year*	23%

Drop out chance
Term 3
33.4%

Development of drop out chance

Prior to enrollment	Term 1	Term 2	Term 3
27.9%	34.6% (6 EC)	24.2% (18 EC)	33.4% (18 EC)

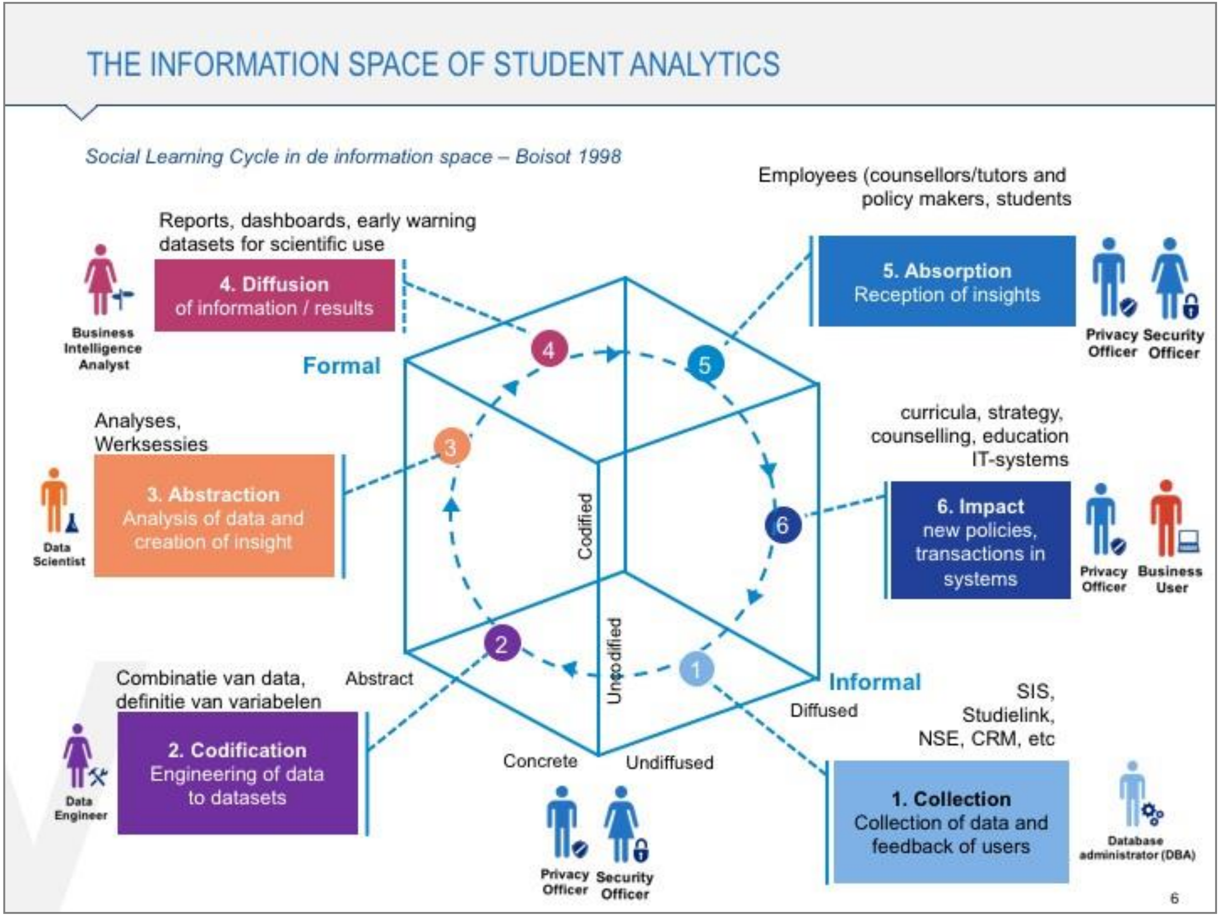
Motivation		*Avg.*	*Top*
• Orientation is present	Yes	92%	92%
• Intensive orientation program	Yes	71%	71%
• Open day	Yes	85%	85%
• Brochure	No	26%	26%
• Confidence in study success	8	8	9
• Proactive study behavior	8	7	8
• Hours paid work and side activities	13	17	7
• Enrolled in two programs at VU	No	5%	5%
• # days since application	237	143	231
• # study programs in the past 2 years	1	1	1
• Avg traveling time by public transport to the VU (minutes)	70	58	19

Performance		*Avg.*	*Top*
• Average grade high school	6.3	6.7	7.5
• Math grade high school	7	6.8	8.1
• Language proficiency grade	75	77	88
• Highest grade	6.5	6.3	8.3
• # EC*	18	12	30
• # EC partial results*	0	0	0
• # no-shows at examinations*	0	0	0

Application date: 08-01-2016
Enrollment date: 09-01-2016
Exit date: 08-31-2017

SAM2016 v10; data update at 02-16-2017 10:30 AM
Green = above average/positive score; orange = average score; red = below average/negative score
* Measure is weighed in the chance to drop out in this term

The personal monitor contained personal information (name, student ID, study program) and all measures that were weighed in the prognosis at any given time in the academic year. The monitor contained dropout prognoses for each study period with the underlying measures. Measures and values were grouped by personal information and general information on the study program, development of the dropout chance, motivational measures and performance measures. A comparison of the student to their average peers and top peers was presented as well. Measures were color coded to express if a score was above average / a positive score, an average score, or below average / a negative score.


THE INFORMATION SPACE OF STUDENT ANALYTICS
Social Learning Cycle in de information space – Boisot 1998
Reports, dashboards, early warning datasets for scientific use
4. Diffusion
of information / results
Business Intelligence Analyst
Formal
Analyses, Werksessies
3. Abstraction
Analysis of data and creation of insight
Data Scientist
Combinatie van data, definitie van variabelen
Abstract
2. Codification
Engineering of data to datasets
Data Engineer
Codified
Uncodified
Concrete
Undiffused
Privacy Officer
Security Officer
Employees (counsellors/tutors and policy makers, students
5. Absorption
Reception of insights
Privacy Officer
Security Officer
curricula, strategy, counselling, education IT-systems
6. Impact
new policies, transactions in systems
Privacy Officer
Business User
Informal
Diffused
SIS, Studielink, NSE, CRM, etc
1. Collection
Collection of data and feedback of users
Database administrator (DBA)
1
2
3
4
5
6

Results: The presentation will focus on the three-year development process of the dataset, the development of the model and the application of the model for student counselling. This process will be tied to the Information Space Theory (Boisot 1998). The results of the student counselling experiment haven't been published yet. Implications: From this case and the three-year process that will be presented, lessons can be learned on the general process of sense making, knowledge making and decision making using student data from transactional information systems within higher education to provide policy makers and student counsellors with evidence based policies and early warning systems.

Resources:

1. Boisot, M.H. (1998), Knowledge Assets – Securing Competitive Advantage in the Information Age, Oxford: Oxford University Press.

DESSI – Ireland’s Data-Enabled Student Success Initiative

Lee O’Farrell
National Forum for the Enhancement of Teaching and Learning in Higher Education (Ireland)

Purpose: Ireland’s National Forum has recently completed a nine-month national project on Learning Analytics, the aims of which were to raise awareness of the many applications of LA for students and institutions, to develop a national community of researchers, practitioners and interested parties and to build capacity across the Irish HE sector to engage with and benefit from this emerging field.

To this end, we are currently developing ORLA (Online Resource for Learning Analytics), an open-access online library of resources that guide institutions and teachers through every step of developing an LA capacity. Resources cover topics such as Data Protection and the GDPR, policy development, ethics, data quality, case studies of lecturers that are using data to inform their module delivery, the LA functionality currently available in VLEs, student communications, effective intervention design etc. These resources will be publicly available from September 2017.

Through the insights developed over the course of this project, we have come to appreciate that LA can only succeed where it is supported by a well-structured, collaborative and sustainable institutional strategy for student success and an institutional culture that recognises the value of data as a resource for supporting learning. This cannot be achieved simply by purchasing an EWS; prediction does not equal prevention.

In order to address this challenge, we have inaugurated DESSI, the national Data-Enabled Student Success Initiative that will initially run from September 2017 to December 2018 to provide support to institutions across Ireland in developing the strategies, implementation plans and cultural revolutions that are necessary to fully harness the benefits of LA in supporting student learning and success.

Design: DESSI is a partnership of representative organisations from across the Irish HE sector, covering Teaching and Learning, IT, Quality Assurance, the national student survey and the representative organisations of our Universities, Institutes of Technology and private Colleges.

In effect, it will provide a publicly-funded consultancy service to our HEI's to support the development of institutionally-tailored strategies and plans in which EWSs are seen as key components of, rather than substitutes for, a holistic approach to enabling student success. We will work with partner institutions to build approaches to this success that are founded on good practice rather than on technology.

One of the key aspects of this initiative is our focus on a multi-disciplinary approach that recognises the necessity of collaborating with stakeholders across the institution. By involving students, lecturers, data owners such as Registry and the Library, data protection and policy officers, educational designers and technologists, senior managers, IT staff, student-support staff, institutional researchers and quality professionals in the process, we hope to develop strategies that are well-planned and informed by good practice at every stage.

In addition to this institutional partnership, we will provide seminars and workshops for all staff to support the development of institutional cultures that recognise the value of data, particularly as a resource for driving student success.

We will also be operating at a sectoral level, exploring the potential for developing shared services and a national architecture to support LA development. This will include an assessment of the current LA platforms, that will enable institutions to identify the products that are most congruent with their needs.

Results: The DESSI initiative is currently in its earliest stages. We do not expect to see any results until the spring semester of 2018 at the earliest. At this point, we expect to have been working with our first tranche of partner institutions for approximately six months.

The results that we expect to begin seeing from this point include:

- The first realisations of collaborative institutional strategies and implementation plans informed by good practice
- An increased awareness of, and commitment to, the use of data to facilitate student success across the Irish HE sector
- The development of individual institutional cultures that recognise the value of data as a resource for supporting students and the importance of comprehensive planning to achieve the maximum impact from the adoption of EWSs and an LA methodology
- An increased number of staff who teach across the Irish HE sector using data to inform their approach to teaching and supporting students
- A willingness on the part of institutions to invest time in developing effective, sustainable strategies before making a financial investment of public funds into analytics platforms.
- The development of a sectoral awareness of the potential for

shared services and a national IT infrastructure

Implications: DESSI is not the world's first nationally-coordinated LA initiative. In fact, DESSI is not an LA initiative *per se*.

We are not aiming to bring LA to the Irish HE sector. We are aiming to support institutions in utilising their data to enhance student success. There is a subtle difference between these two approaches that is embodied by how we measure the initiative's success. An LA-focused initiative can be assessed quantitatively through the growth in use of early warning systems etc.

DESSI's focus, however, is on strategy, good practice and institutional culture, rather than on technology. We hope to develop a national understanding that developing an LA capability is an essential component for supporting student success. It is not an end in itself.

If we are successful in our endeavours, we hope to overcome many of the major hurdles to implementing an effective data-informed strategy for student success, not in a single institution but across an entire national sector.

Learning analytics dashboard for improving the course passing rate in a randomized controlled experiment

Jan Hellings

Amsterdam University of Applied Science

Purpose: A Learning analytics dashboard (LAD) (Figure 1) is developed to increase the success rate of the Java programming course by encouraging the online activities of the students in the Learning Management Systems (LMS) Moodle1 and MyProgrammingLab2 (MPL). The study is carried out among freshman computer science students of the Amsterdam University of Applied Science who attend an eight-week Java programming course to acquire their basic Java programming skills. The Java programming course is set up like a blended learning course (Bos, 2016, p. 12). The LAD is designed to visualize online learning behavior and help students to improve self-knowledge by reviewing and analyzing their personal online history (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013), by predicting their result based on their online behavior and comparing their individual result with the average cohort results. The study addresses the following research questions:

1. What is the effect of the learning analytics dashboard on the passing rate and the grades of the students participating in the Java programming course?
2. What is the effect of the learning analytics dashboard on the online activities of students?

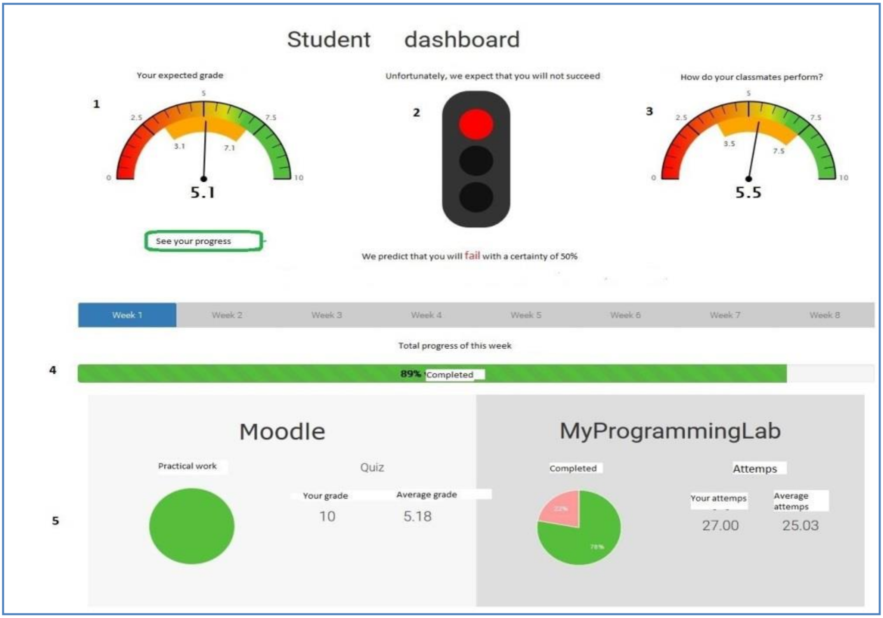


Figure 1 LAD of student in week 1

Design: The study is setup as a Randomized Controlled Trial (RCT) (Figure 2). A total of 556 students are involved in the experiment. 276 in the treatment group receiving a dashboard and 280 in the control group receiving no dashboard.

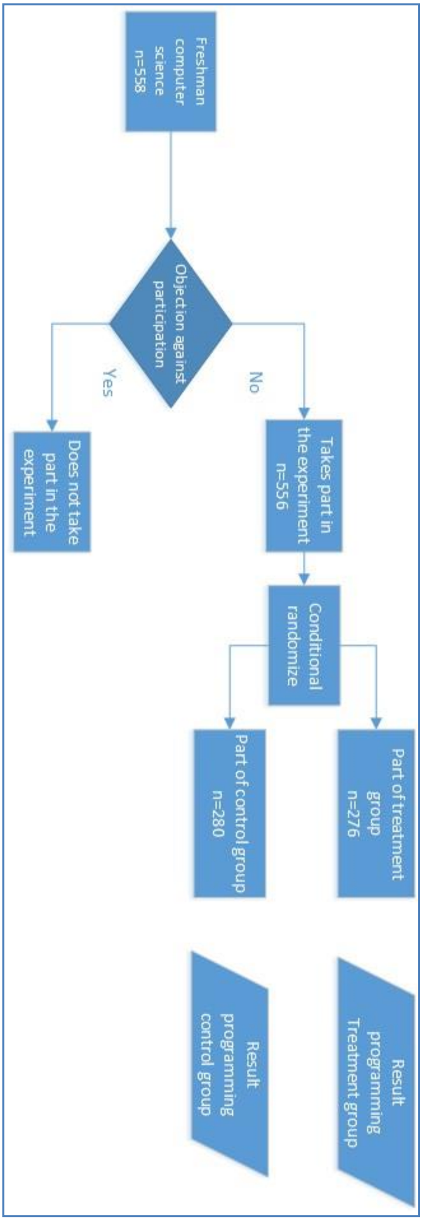


Figure 2 Experimental setup for the dashboard intervention

Results: The effect of the dashboard treatment on passing the programming course is shown in Table 1, where the means of passing the exams for the treatment- and control groups are shown. All exam is the summation of first and retake exam. This shows that for the treatment group the average is slightly higher in the percentage passing at the first exam. However, this difference is not statistically significant.

Table 1 The mean of passing the exams of the treatment- and control group

Result	Control	Dashboard	n	missing	p
First exam passed	.523	.533	228	328	.830
Retake passed	.155	.519	116	440	.599
All exam passed	.741	.714	456	100	.516

n=556. Passed: 332 Failed= 124

The analysis of the online behaviour showed no significant difference in the online behaviour between the control – and treatment group.

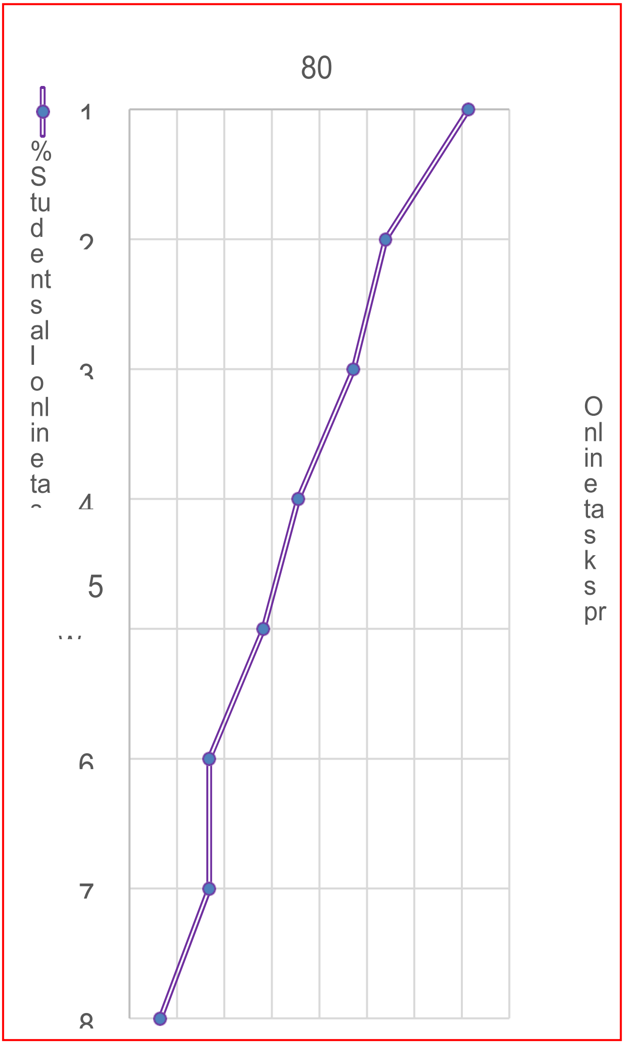


Figure 3 Percentage of students completed all online tasks per week.

Furthermore, the analysis of the online behavior showed that online activity strongly declined over time (Figure 3). The overall online usage dropped from about 70% in week 1 to about less than 50% in week 3. After two weeks students probably figured out that there were no sanctions against poor online performance and they stopped using Moodle and MPL.

A questionnaire was send out to investigate the student's opinions of the dashboard. It showed that the students liked the overview capabilities of the dashboard but did not like the prediction models. In sum, it is likely that students did not use the dashboard that often, because the prediction models were not predicting very well. The inaccurate prediction is probably due to the drop of online activities of the students during the course (Figure 3).

Implications: In future research is we will take the following aspects into account. The use of prediction models in a blended learning course with only online formative assessments is quite cumbersome, because students tend to do only their online exercises if there is something to gain. The prediction models cannot predict if there is no online activity. The prediction models could be used in blended learning courses with online summative assessments (Tempelaar, Rienties, & Giesbers, 2015) or fully online courses (Hu, Lo, & Shih, 2014). The survey conducted among students who used the dashboard showed that they appreciated the overview of the LAD but did not like the prediction models and the strict tone of the mails. In future research, it might be wise to only shows their online progress, without the prediction models and use more persuasion in the form of a game (Llagostera, 2012) to encourage the students to do their online tasks.

Resources:

1. Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, *36*(0), 469–478. http://doi.org/http://dx.doi.org/10.1016/j.chb.2014.04.002
2. Llagostera, E. (2012). On Gamification and Persuasion. *Proceedings of SBGames 2012*, 12– 21.
3. Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, *47*, 157–167. http://doi.org/10.1016/j.chb.2014.05.038
4. Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning Analytics Dashboard Applications. *American Behavioral Scientist*, *57*(10), 1500–1509. http://doi.org/10.1177/0002764213479363

Session 3 - Learning Analytics infrastructure and dashboard

Designing a Learning Analytics Capability Model

Justian Knobbout
HU University of Applied Sciences Utrecht

Purpose: To perform all activities required to turn student data into actionable insights, the development, implementation and deployment of learning analytics systems is required. Nevertheless, a vast amount of researchers primarily focusses on the development phase (cf. Macfadyen & Dawson, 2010; Romero-Zaldivar et al., 2012; Baker et al., 2015). That is, what is technically possible, which measures should be included, what requirements do systems have or in what way can these requirements designed into the system. There are only few documented large-scale deployments of fully functional learning analytics systems (Ferguson et al., 2014). Examples of these systems are Course Signals (Arnold & Pistilli, 2012), RioPACE (Smith, Lange & Huston, 2012), and Open Academic Analytics Initiative (Jayaprakash et al., 2014; Lauría et al., 2013). Many other institutes are curious to discover the benefits of learning analytics as well and are exploring ways to implement it in their operations. This is not an easy exercise, as many challenges arise, e.g., changes to the existing information systems by implementing a Learning Record Store and customizing data streams (Apereo, 2015; Del Blanco et al., 2013); managing the increase in workload for teachers (Whale, Valenzuela & Fisher, 2013); and making sure to comply with privacy legislation (Jisc, 2015). In 2015, SURFnet released a report on the challenges of learning analytics for Dutch educational institutes (Berg et al., 2015). It shows that learning analytics has a lot of potential but is still limitedly embedded in digital learning environments. Trigt (2016) describes the pedagogical, ethical, legal, and technical issues experienced by institutes when applying learning analytics. Existing models often fail to overcome these challenges and issues (Colvin et al., 2015). Whilst some of mentioned problems relate to either existing or to- be obtained assets, others involve the necessary capabilities to effectively use these assets.

Challenges like these lie at the heart of the resource-based view. The resource-based view (RBV) attributes an organization's performance to its resources (assets and capabilities) (Bharadwaj, 2000). Whilst some of the challenges relate to either existing or to-be obtained assets, others involve developing the necessary capabilities to effectively use these assets. This PhD study draws from the resource-based view and aims at providing an answer to the following main research question: *"How can learning analytics benefit teachers and learners in Dutch higher educational institutes?"*

Design: Our perspective will primarily be from an Information Systems (IS) point of view as many of the drivers of learning analytics involve IS aspects like data capturing and storing, tool availability, and digitalization of education (Ferguson, 2012; Baker & Siemens, 2014). A design science approach for IS research as elaborated by Hevner et al. (2004) is proposed. To provide an answer to the sub questions and, consequently, the research main question, five studies will be conducted (see Table 1).

	Study 1: Determining dependent variables	*Study 2:* Learning analytics capabilities (literature)	*Study 3:* Learning analytics capabilities (case studies)	*Study 4:* Learning Analytics Capability Model validation (focus groups)	*Study 5:* Learning Analytics Capability Model validation (pilot project)
Sub question 1: Beneficial learning analytics	**X**				
Sub question 2: Learning analytics capabilities		**X**	**X**	**X**	
Sub question 3: Effect of capabilities					**X**

Table 1: Studies in Relation to Research Sub Questions

Study 1 tries to determine what measures of successful learning analytics are. That is, what are the variables one tries to influence when performing learning analytics activities? A literature review of extant learning analytics literature is conducted to formulate an answer on this question. Study 2 involves a literature review as well, in which we analyse existing literature from both the learning analytics field and other domains to identify already mentioned capabilities (*processes, capacities)* necessary for organizational-broad data analytics. In study 3 we will find more of those capabilities during case studies in both the educational and non- educational domain. From the results of studies 1 to 3, a Learning Analytics Capability Model is designed. This model will be validated by means of focus groups with experts (study 4) and a large-scale pilot project (study 5).

Results: Studies 1 and 2 are momentarily conducted and will be finished end of October 2017. Primarily results of study 1 show that few research articles involve complete learning analytics processes including interventions but those who do provide a variety of intended measurable outcomes, e.g., improved course quality, enhanced communication skills, and higher grades. Primarily results of study 2 show that some generic models on learning analytics implementation exist but these models are quite abstract and are hard to translate into practice. Moreover, these models (often) look at implementing learning analytics but not sustainable application of it and what assets and capabilities are required in order to do so.

Implications: Existing learning analytics models and frameworks often fail to translate theory into practice (Colvin et al., 2015). This PhD research aims at contributing to the knowledge base of the learning analytics field. That is, identifying what capabilities should be built by higher educational institutes in order to successfully apply learning analytics. The research adds prescriptive knowledge by means of exaptation (Gregor & Hevner, 2013) and validates it in practice by means of focus groups and a pilot project. Few good examples of large-scale learning analytics system application do currently exist (Ferguson et al., 2014) so extended research in this area will have academic relevance as it elaborates on a gap in knowledge.

Resources:

1. Apereo (2015) The Learning Analytics Diamond.
2. Arnold, K. E., & Pistilli, M. D. (2012). Course signals at purdue: Using learning analytics to increase student success. Paper presented at the *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge,* 267-270.
3. Baker, R. S., Lindrum, D., Lindrum, M. J., & Perkowski, D. (2015). Analyzing early at-risk factors in higher education E-learning courses. *International Educational Data Mining Society,*
4. Berg, A., Bogaard, M. van den., Drachsler, H., Filius, R., Manderveld, J., & Schuwer, R. (2015). Grand challenges learning analytics & open en online onderwijs: een verkenning [whitepaper]. Retrieved from https://www.surf.nl/binaries/content/assets/surf/nl/kennisbank/2015/rapport-grand- challenges-la--ooo---een-verkenning.pdf on 13 September 2016.
5. Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly,* , 169-196.
6. Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S., & Fisher, J. (2015). Student retention and learning analytics: A snapshot of australian practices and a framework for advancement. *Sydney: Australian*

Office for Learning and Teaching,

7. Del Blanco, Á., Serrano, Á., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2013). E- learning standards and learning analytics. can data collection be improved by using standard data models? Paper presented at the *Global Engineering Education Conference (EDUCON), 2013 IEEE,* 1255-1261.
8. Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5-6), 304-317.
9. Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). Setting learning analytics in context: Overcoming the barriers to large-scale adoption. Paper presented at the *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge,* 251-253.
10. Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly, 37*(2), 337-355.
11. Hevner,A.,March,S.,Park,J.,&Ram,S.(2004).Designsciencein informationsystems research. *MIS quarterly*, *28*(1), 75-105.
12. Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). An open source analytics initiative.
13. Lauría, E. J., Moody, E. W., Jayaprakash, S. M., Jonnalagadda, N., & Baron, J. D. (2013). Open academic analytics initiative: Initial research findings. Paper presented at the *Proceedings of the Third International Conference on Learning Analytics and Knowledge,* 150-154.
14. Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers & Education, 54*(2), 588-599.
15. Romero-Zaldivar, V., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education, 58*(4), 1058-1067.

16. Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61.
17. Trigt, M. van. (2016). Hoe data de kwaliteit van het hoger onderwijs kunnen verbeteren [whitepaper]. Retrieved from https://www.surf.nl/binaries/content/assets/surf/nl/kennisbank/2016/whitepaper-la-web- def.pdf on 13 September 2016.
18. Whale, S., Valenzuela, F., & Fisher, J. (2013). Implementing timely interventions to improve students' learning experience. *Electric Dreams.Proceedings Ascilite,* , 908

The Apereo Learning Analytics Initiative. How innovation, community building and 100% Open works on the Global Stage.

Ian Dolphin, Alan Berg, Patrick Lynch
Apereo Foundation

Purpose: In this presentation, we will discuss the Apereo Learning Analytics Initiative, what it is, a bit about how we implemented Learning Analytics and deployed at a national level and a discussion on how many Universities work together to decrease risk and improve a quality of the students and teachers analytics experience. We will also consider the future priorities of the initiative.

The Apereo Foundation is an open source foundation serving Education (AF, 2015). The foundation has a global presence with many hundreds of Universities using their products such as Yale CAS (an SSO solution), uPortal, Apereo OAI, rostering systems and Sakai, a well-known Learning platform (Berg & Dolphin, 2011), etc. (AF, 2017).

In a period of uncertainty around the failure of large scale projects associated with sharing data, such as the Inbloom experience (Kharif, 2014) and the shifting sands of European legal requirements (Hoel, Griffiths & Chen, 2017), the Apereo foundation and its co-development 100% open community support, represents a safe partnership to discuss and build on common requirements.

Design: The Apereo Learning Analytics Initiative infrastructure contains three elements: Collection and storage of data is provided by the Open Learning Record Warehouse, that collects Student Activity streams via the xAPI standard (Berg et al, 2016): Visualisation and reporting are provided by a highly extensible Open Dashboard and interventions and case management is handled by the Students Success Plan. Most of the infrastructure is deployed as part of the JISC, UK experimental LA infrastructure (Sclater., Berg & Webb, 2015). Other open Standards being applied include IMS Caliper and IMS OneRoster and PMML (an XML format to share Machine Learning Algorithms). Furthermore, the OpenDashboard is a web application which is accessible via IMS LTI.

Results: The results are large scale collaborations and deployments, 100% Open communities, experimentation via data hackathons at conferences based on the infrastructure (Cooper et al, 2017) and support for wider debates.

The Apereo Learning Analytics Initiative has rock solid foundations within research (Jayaprakash et al, 2016; Jayaprakash et al, 2014) and experiences with integrating third party tools such as SNAPP, a tool for visualizing Social Networking Activity (Bakharia & Dawson, 2011), into its Learning Management System Sakai. Experience has shown that the models for estimating student risk are stable across organizations and the community adapt quickly to change, for example, an earlier element, the Learning Analytics Processor, has been augmented by the Hadoop ecosystem which uses XML scripts as the implementation model. The aim is that the scripts already developed for specific instances will be made available under an appropriate open license soon to act as the reference for the community.

Increasingly, the Apereo community is turning attention to the barriers to wider adoption and broader sources of data. For example, we heavily rely on standards to provide for interoperability, however, also supports local variation within the framework. In practice, this means that models can be taken from the core functionality and then modified to deal with local requirements e.g. shifting focus from students failing to complete to students on the boundary between grading bands. This experimentation is enabled by the standards applied and supported through a community of motivated collaborators.

We will conclude the presentation with a summary of “lessons learned” from the Apereo learning analytics work to date, and a forward look at emerging priorities.

Implications: At a practical level co-development of an 100% Open Learning Analytics Initiative ensures a high degree of openness that stimulates competition and transparent improvement in the LA eco sphere.

Acknowledgments: We would like to acknowledge the work of all those dedicated and spirited members within the Apereo communities who make the online learning experience of many millions of students possible.

Resources:

1. AF. (2015) Apereo Foundation core values. Retrieved from https://www.apereo.org/sites/default/files/projects/Brochures/Apereo%20Brochure%2018OC T15.pdf
2. AF. (2017) The Apereo Foundations communities and software range. Retrieved from https://www.apereo.org/content/projects-communities
3. Bakharia, A., & Dawson, S. (2011). SNAPP. In Proceedings of the 1st International Conference on Learning Analytics and Knowledge - LAK '11 (p. 168). New York, New York, USA: ACM Press. http://doi.org/10.1145/2090116.2090144
4. Berg, A., & Dolphin, I. (2011). *Sakai CLE Courseware Management: The Official Guide*. Packt Publishing Ltd.
5. Berg, A., Scheffel, M., Drachsler, H., Ternier, S., & Specht, M. (2016). The dutch xAPI experience. In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16 (pp. 544–545). New York, New York, USA: ACM Press. http://doi.org/10.1145/2883851.2883968
6. Cooper, A., Berg, A., Sclater, N., Dorey-Elias, T., & Kitto, K. (2017). LAK17 hackathon. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17 (pp. 514–515). New York, New York, USA: ACM Press. http://doi.org/10.1145/3027385.3029435
7. Hoel, T., Griffiths, D., & Chen, W. (2017). The influence of data protection and privacy frameworks on the design of learning analytics systems. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17 (pp. 243–252). New York, New York, USA: ACM Press. http://doi.org/10.1145/3027385.3027414
8. Jayaprakash, S. M., Laur'\ia, E. J. M., Gandhi, P., & Mendhe, D. (2016). Benchmarking Student Performance and Engagement in an Early Alert Predictive System Using Interactive Radar Charts. Proceedings of the Sixth

International Conference on Learning Analytics & Knowledge, 526–527. http://doi.org/10.1145/2883851.2883940

9. Jayaprakash, S. M., Moody, E. W., Lauria, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. Journal of Learning Analytics, 1(1), 6–47. http://doi.org/10.18608/jla.2014.11.3
10. Kharif, O. (2014). Privacy Fears Over Student Data Tracking Lead to InBloom's Shutdown
11. Businessweek. Retrieved from
12. http://www.bloomberg.com/bw/articles/2014-05-01/inbloom-shuts-down-amid-privacy-fears- over-student-data-tracking
13. Sclater, N., Berg, A., & Webb, M. (2015). Developing an open architecture for learning analytics. Proceedings of the EUNIS 2015 Congress. http://doi.org/ISSN: 2409-1340

Learning analytics and lecturer's knowledge on learning analytics dashboards

Nils Siemens MSc

Amsterdam University of Applied Sciences (AUAS/HvA)

Purpose: This presentation discusses the intersection between Technology, Pedagogy And Content Knowledge (TPACK) and Learning Analytics. TPACK conceptualizes the mentioned knowledge fields, that lecturers need to apply to technology to enhance effectively education (Voogt, Fisser, Pareja Roblin, Tondeur, & van Braak, 2013). TPACK is about the 'nuanced understanding' of 'the complex relationships' between the three knowledge fields that made up TPACK (Mishra & Koehler, 2006).

The risk for all educational technologies is the unbalanced application whose effect can be mitigated by 'adequate staff training and support' (Rienties et al., 2013). For learning analytics dashboards, for example for feedback, unbalanced use of dashboards as a feedback instrument is a known risk as described by (Gašević, Dawson, & Siemens, 2015).

Unfortunately, Jivet, Scheffel, Drachsler & Specht, 2017 found that a minority of dashboards are grounded in educational stances/theories such as Cognitivism and Educational Design Verbert et al (2014) argued that learning analytics dashboards support learning and teaching on different levels: promotion of awareness, reflection, sense making, and impact on behavior or the generation of new meanings. The author will argue that for the betterment of Learning Analytics dashboards lecturers require knowledge on TPACK. The question is which combination of knowledge is most desirable for lecturers to use learning analytics dashboards in an effective and efficient way. To name a few combinations: Technological knowledge on data representation, pedagogical knowledge on how to discuss dashboard-results with students, the educational conceptual background of the dashboard, content knowledge about the subject taught in relation to the dashboard result.

This presentation will shine a light on the usefulness of the TPACK knowledge framework for lecturers while they are designing or choosing or deploying learning analytics dashboards. The lecturers considering the early stages of development of dashboards, the technological complexities, and variation in educational concepts behind dashboards and the different levels of knowledge held by lecturers about the technological specific's.

The following research question is proposed: RQ1: Which TPACK knowledge on learning analytics dashboards are found in the literature? RQ2: Which TPACK knowledge can be stated as a baseline for lecturers, considering the technological complexities, early stages in development and variation in educational concepts behind dashboards?

Design: This study consists of a literature review about TPACK and learning analytics dashboards knowledge as is necessary for lecturers. The literature review will highlight connections between the fields of learning analytics dashboards- and TPACK lecturer knowledge literature. These findings are discussed in the session.

Results: The literature review results are used in the session as a jump off point for the formulation of research directions. The result of the session is a map of directions based on the discussion and opinion of attendees grounded in the theory presented.

Implications: The results of the literature research, the formulated opinion and discussion contribute to the insights on how TPACK can be used as a framework to guide lecturer's knowledge development on learning analytics and will help to identify the pitfalls and possibilities in the implementation of learning analytics dashboards in lecturer's daily work.

Acknowledgments: I would like to address special thanks to my department (Education and Research, ICT-Services) and my colleagues for supporting this proposal and successful completion.

Resources:

1. Gašević, D., Dawson, S., & Siemens, G. (2015). Let’s not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71.
2. Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough. Pitfalls of learning analytics dashboards in the educational practice. Retrieved from http://dspace.ou.nl/handle/1820/7985
3. Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *The Teachers College Record*, *108*(6), 1017–1054.
4. Rienties, B., Brouwer, N., Bohle Carbonell, K., Townsend, D., Rozendal, A.-P., van der Loo, J., … Lygo-Baker, S. (2013). Online training of TPACK skills of higher education scholars: a cross-institutional impact study. *European Journal of Teacher Education*, 1–16. https://doi.org/10.1080/02619768.2013.801073
5. Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*(6), 1499–1514.
6. Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., & van Braak, J. (2013). Technological pedagogical content knowledge - a review of the literature: Technological pedagogical content. *Journal of Computer Assisted Learning*, *29*(2), 109–121. https://doi.org/10.1111/j.1365-2729.2012.

Organising and Sponsoring Partners

Amsterdam Center for Learning Analytics

eduworks

VU VRIJE UNIVERSITEIT AMSTERDAM

Faculteit der Gedrags- en Bewegings- wetenschappen

UNIVERSITEIT VAN AMSTERDAM

SURF