Editorial

doi: 10.1111/jcal.12139

Journal of Computer Assisted Learning

Learning analytics in massively multi-user virtual environments and courses

There is much ongoing interest in big data and the role it can play in decision-making in diverse areas of science, commerce and entertainment. By employing a combination of modern artificial intelligence, machine learning and statistics techniques, extremely large and complex data sets can be 'mined' in a variety of ways to reveal relationships, patterns and insights not easily discoverable through standard database management tools and data processing applications. In education, data mining approaches have been applied to the analysis of electronic stores or repositories of student data for a number of years now (Romero & Ventura, 2007), but this has been occurring largely at the institutional or sector level. Such applications, which are sometimes referred to as 'academic analytics' (Campbell, DeBlois, & Oblinger 2007; Goldstein & Katz, 2005), have not become mainstream, being relevant mainly to governments, funding agencies and institutional administrators rather than students and teachers (Siemens *et al*., 2011). More recently, a new field known as *learning analytics* (Long & Siemens, 2011; Siemens *et al*., 2011) has emerged that seeks to generate knowledge 'about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs' (Siemens, 2011, para. 5). This knowledge can be employed for a range of purposes, among which are to allow learners to reflect on their activity and progress in relation to that of others as well as to assist teachers and support staff in predicting, identifying and supporting learners who may require additional attention and intervention (Powell & MacNeill, 2012).

Occurring in parallel is the burgeoning trend towards the delivery of education and learning at a 'massive' scale. The last decade has seen an explosion of activity in the use of massively multiplayer online games (e.g. World of Warcraft) and virtual worlds (e.g. Second Life) for both formal and informal learning (Childress & Braswell, 2006; Dalgarno & Lee, 2010). These *massively multi-user virtual environments* (MMVEs) are rife with opportunities for exploiting learning analytics methods to produce enhanced outcomes and experiences for students. At the same time, we have been witnessing a movement in which many universities and colleges, including some of the most prestigious institutions of higher learning in the world (e.g. Harvard, Stanford, MIT and the Universities of Melbourne, Toronto and Edinburgh, to name a few), are 'opening up' their course offerings to massive numbers of participants on the Internet (see, for example, Brown, 2013; Daniel, 2012; Jona & Naidu, 2014; McAuley, Stewart, Siemens & Cormier, 2010; Siemens, Irvine, & Code, 2013). In such *massive open online courses* (MOOCs), the involvement of hundreds, thousands or even tens of thousands of students creates a heighted imperative to formulate alternative strategies for feedback and assessment that are less reliant on individual teachers. Learning analytics have the potential to be used in MOOCs to facilitate new models of self and peer assessment as well as to make possible the implementation of automated mechanisms to support and augment students' self-regulated learning goals and processes.

This special issue of JCAL addresses the intersection of learning analytics on one hand and MOOCs and MMVEs on the other, its primary goal being to help foster and encourage the interdisciplinary dialogue and exchange needed to bring together the various contributory bodies of knowledge encompassed by the two domains. The six articles contained within the issue individually and collectively highlight both the predictive and prescriptive capabilities of learning analytics as applied to 'massive' situations, demonstrating how they can be harnessed in different ways to assist us in better understanding, and thus better serving, learners and learning.

The special issue opens with an article by Saif Rayyan, Colin Fredericks, Kimberly Colvin, Alwina Liu, Raluca Teodorescu, Analia Barrantes, Andrew Pawl, Daniel Seaton and David Pritchard, who present a case study of an introductory physics MOOC based on blended pedagogy that evolved from materials originally created for a face-to-face course. The authors show how learning analytics were used to understand the impact of various elements of the MOOC course design on student behaviour and to inform iterative development and refinement of the design, which ultimately led to an increase in retention rates.

The second and third articles hone in on a key issue in MOOC arena: student motivation. Bart Pursel, Liang Zhang, Kathryn Jablokow, Gi Woong (Josh) Choi and Darrell Velegol's research systematically examined student demographic data, intended behaviours and course interactions in an effort to identify predictors of MOOC completion, while Paula de Barba, Gregor Kennedy and Mary Ainley looked specifically at the effect of motivation and participation on students' performance in a MOOC, focusing in particular on those students persisting to the end of the course. The outcomes and findings of these two studies have implications for the use of learning analytics to provide formative feedback and support to students as well as to assist in adapting course design and delivery to maximize student success.

The next two articles in the special issue are concerned with learning analytics as they relate to the social aspects of MOOCs. In the fourth article, Carlos Alario-Hoyos, Pedro Muñoz-Merino, Mar Pérez-Sanagustín, Carlos Delgado Kloos and Hugo Parada G. describe how they used data drawn from five social tools in a MOOC to characterize the top contributors and identify variables that may aid in identifying those students early in the course. They found a moderate positive correlation between contribution level (measured in terms of number of posts) and performance (measured in terms of final scores), and also considered the roles played by top contributors in assisting their peers, for example, by assuming partial responsibility for tasks traditionally performed by the teacher.

Sean Goggins, Krista Galyen and James Laffey, the authors of the fifth article, conducted a mixed-methods exploratory study aimed at linking social learning structure with performance in a MOOC that used a curriculum designed for small group work. They devised a novel multi-dimensional performance construct along with an innovative approach for modelling the social structure of MOOC participants and for connecting the social structure and the performance measures. Though Goggins *et al*. do not attempt to make general claims of causality or correlation between the structural properties of groups and the level at which the members of those groups perform, their study and its findings do illustrate new ways of viewing the relationship between group structure and performance, pointing to types of learning analytics that may prove useful in the future and that are unlikely to be uncovered through computational analysis alone.

In the sixth and final article, by Ryan Baker, Jody Clarke-Midura and Jaclyn Ocumpaugh, the learning environment in question is an avatar-based multi-user virtual environment rather than a MOOC. Using log file data from almost 2000 middle school students, the authors developed models of user interaction within the environment for predicting whether a student will successfully complete a scientific inquiry task. They accomplished this by identifying behaviours that lead a student to discriminate between causal and non-causal factors, enabling the student to draw a correct final conclusion and to craft a causal explanation for the conclusion. Baker *et al*. provide in their article a detailed example of how their models can be easily adapted from one virtual scientific inquiry scenario to another.

At the end of the issue is an invited epilogue in which Hendrik Drachsler and Marco Kalz reflect on the current state of play and ongoing challenges facing research in the area of learning analytics in massively multi-user environments and courses, as manifested in the six special issue articles. They propose a framework for conceptualizing innovations in the area, then attempt to classify and analyze the studies and initiatives reported in the special issue articles with respect to their framework. They conclude with a discussion of a number of aspects in which they feel additional work and developments are prudent.

It is hoped that this special issue will be a useful knowledge base and source of information pertaining to how learning analytics can be used for a variety of research and practical applications involving massive numbers of students, and that its contents will act as a catalyst for further discourse and studies in this still-nascent but very important and promising area of scholarship.

M.J.W. Lee,* P.A. Kirschner,† & L. Kester‡

School of Education, Charles Sturt University, Australia
†*Welten Institute, Open University of the Netherlands, The Netherlands*
‡*Department of Pedagogical and Educational Sciences, Utrecht University, The Netherlands*

References

Brown, J. (Ed.) (2013). MOOCs and technology [Special issue]. *Research and Practice in Assessment*, *8*. Retrieved from http://www.rpajournal.com/dev/wp-content/uploads/2013/05/Vol8_Summer.pdf

Campbell, J. P., DeBlois, P. B., & Oblinger, D. (2007). Academic analytics: a new tool for a new era. *EDUCAUSE Review*, *42*, 41–57.

Childress, M. D., & Braswell, R. (2006). Using massively multiplayer online role-playing games for online learning. *Distance Education*, *27*, 187–196. doi:10.1080/01587910600789522.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *40*, 10–32. doi:10.1111/j.1467-8535.2009.01038.x.

Daniel, J. (2012). Making sense of MOOCs: musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, *3*. doi:10.5334/2012-18.

Goldstein, P. J., & Katz, R. N. (2005). Academic analytics: the uses of management information and technology in higher education. Retrieved from http://www.educause.edu/ir/library/pdf/ers0508/rs/ers0508w.pdf

Jona, K., & Naidu, S. (2014). MOOCs: emerging research [Special issue]. *Distance Education*, *35*. doi:10.1080/01587919.2014.928970.

Long, P., & Siemens, G. (2011). Penetrating the fog: analytics in learning and education. *EDUCAUSE Review*, *46*, 31–40.

McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC model for digital practice. Retrieved from http://www.elearnspace.org/Articles/MOOC_Final.pdf

Powell, S., & MacNeill, S. (2012). Institutional readiness for analytics. Retrieved from http://publications.cetis.ac.uk/wp-content/uploads/2012/12/Institutional-Readiness-for-Analytics-Vol1-No8.pdf

Romero, C., & Ventura, S. (2007). Educational data mining: a survey from 1995 to 2005. *Expert Systems with Applications*, *33*, 135–146. doi:10.1016/j.eswa.2006.04.005.

Siemens, G. (2011). Call for papers for the first international conference on learning analytics and knowledge (LAK'11). Retrieved from https://tekri.athabascau.ca/analytics/

Siemens, G., Irvine, V., & Code, J. (Eds.) (2013). Massive open online courses [Special issue]. *MERLOT Journal of Online Learning and Teaching*, *9*. Retrieved from http://jolt.merlot.org/Vol9_No2.htm

Siemens, G., Gašević, D., Haythornthwaite, C., Dawson, S., Buckingham Shum, S., Ferguson, R., Duval, E., Verbert, K., & Baker, R. S. J. d. (2011). Open learning analytics: an integrated and modularized platform. Retrieved from http://www.solaresearch.org/OpenLearningAnalytics.pdf