HERDSA Review of Higher Education, Vol. 6, 2019
www.herdsa.org.au/herdsa-review-higher-education-vol-6/37-63

The Journey of Learning Analytics

Srećko Joksimović*, Vitomir Kovanović and Shane Dawson
University of South Australia, Adelaide, Australia

Abstract

It has been almost a decade since the emergence of learning analytics, a bricolage field of research and practice that focuses on understanding and optimising learning and learning environments. Since the initial efforts to make sense of large learning-related datasets, learning analytics has come a long way in developing sophisticated methods for capturing various proxies of learning. Researchers in the field also quickly recognised the necessity to tackle complex and often controversial issues of privacy and ethics when dealing with learner-generated data. Finally, despite huge interests in analytics across various stakeholders—governments, educational institutions, teachers, and learners—learning analytics is still facing many challenges when it comes to broader adoption. This article provides an overview of this journey, critically reflecting on the existing research, providing insights into the recent advances, and discussing the future of the field, positioning learning analytics within the broader agenda of systems thinking as means of advancing its institutional adoption.

Keywords: learning analytics; higher education; analytics; technology enhanced learning.

1. Introduction

Numerous industries such as health, banking, insurance, aviation, entertainment and telecommunications have long seen the advantages in leveraging the insights brought about by the analysis of large-scale data (Kiron, Shockley, Kruschwitz, Finch, & Haydock, 2012; Manyika et al., 2011; Siemens, 2013). From optimised flight paths to predictive health insurance models, the use of big data has disrupted industries and transformed consumer behaviour. In almost stark contrast, the education sector has

* Email: Srecko.Joksimovic@unisa.edu.au

been slow, or at best cautious, in terms of utilising the vast array of data generated and collected during student learning. The lack of data utilisation is surprising, given that educational technologies such as the learning management systems (LMS) are well established and mostly considered as a core resource for contemporary teaching practice. However, it is only relatively recently that education organisations have begun to dip into the very deep waters of data analytics and machine learning to provide insights into teaching quality and student learning experiences.

In early 2011, a small group of educational researchers hosted *The First International Learning Analytics (LAK'11) Conference* in Banff, Canada. A goal of this first gathering was to define and scope the emergent research focusing on understanding student learning through the use of machine learning, data mining and data visualisation methods. Outcomes from this initial conference included the formation of the Society for Learning Analytics Research (SoLAR)[1] and the defining of learning analytics as the *"measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs"* (Siemens, Long, Gašević, & Conole, 2011, para. 4). From this small gathering, the field has witnessed a dramatic uptake in interest through research funding, publications and commercialisation of associated technologies. The interest in learning analytics stems from the field's connection to the use of technologies in education alongside its perceived benefits in addressing the challenges often associated with contemporary teaching practice. For example, learning analytics can assist in providing personalised and timely assessment and feedback at scale to large-sized classes.

Modern education institutions are required to balance their role as a public good alongside the need to remain financially viable–if not profitable. Early work in learning analytics was seen to provide solutions to balance this agenda. Increased student diversity and demand in a context of reduced government funding called for more cost-effective models of education while maintaining high levels of teaching quality. In simplistic terms these drivers resulted in increased class sizes and wide-spread adoption of learning technologies to promote more flexible access to education. Learning analytics uses the available student learning data in such naturalistic settings to establish early indicators of student attrition and academic performance. Clearly, the capacity to provide early interventions to retain students had a direct financial incentive for institutions that also provided a public good. In this instance, the potential for improved student learning experiences through timely feedback and support. Yet, while learning analytics is framed as a new field of research, the concept of analysing data about learners and their contexts is not new to education (see Figure 1). The utilisation of various forms of learning technologies and learning data were known to educational research long before the emergence of learning analytics. Much of the data mining techniques and methods now commonly used in learning analytics research, such as social network or discourse analysis, have a long history outside of education (Baker & Yacef, 2009; Dawson, Gašević, Siemens, & Joksimovic, 2014). However, the establishment of learning analytics as a discrete field has served to act as a catalyst for coalescing multiple research domains, methods and theories of learning to provide new opportunities of investigation for understanding the learning process.

This review provides a historical overview of the development of learning analytics from the genesis of contributing fields of work, through to early forays into predictive models of student performance, to the more recent generation of fine-grained insights into learning processes. In so doing, the review frames learning analytics as a field that is firmly rooted in both social and technical research ideologies. This duality brings a high degree of complexity as well as potential to transform education practice– particularly when considering the range of applications of learning analytics research. The following section outlines the past research and future directions in learning analytics, noting the transitions from a field focused on student retention to more sophisticated analysis of learning processes. The review outlines the shift from individualised analyses towards more group and social-based practices. Accordingly, the data sources employed in learning analytics research have evolved from single sources of student learning data (e.g. LMS) to multimodal, integrating multiple data sources. Future areas of investigation are discussed including the challenges and opportunities for research and practice.

2. Learning analytics as a field of research

Learning analytics is considered a bricolage field. That is, a field of research that spans multiple, yet well-established disciplines (Gašević, Kovanović, & Joksimović, 2017). Learning analytics draws on theories and methods from machine learning and data science, education, cognitive psychology, statistics, computer science, neuroscience, and social and learning sciences to name but a few (Baker & Inventado, 2014; Doug Clow, 2013; Siemens, 2013). Although learning analytics is frequently touted as an emerging field of research and practice, it does build on a rich history of related disciplines (see Figure 1) that establish the basis for learning analytics inquiries (Dawson, Joksimović, Poquet, & Siemens, 2019; Reimann, 2016).

However, learning analytics does differ from more traditional education analyses in a number of ways. Firstly, due to its strong quantitative focus, the size of data sets tends to be significantly larger allowing for a greater level of confidence in the generalizability of the findings (Reimann, 2016). Secondly, as data is mostly collected from technical systems, there is a very fine level of granularity of available variables that cannot be captured through observational studies, interviews or self-reports (Reimann, 2016). Lastly, the data tends to be longitudinal. That is, the manner of data collected and the processes used for collecting provide for a strong temporal dimensionality to be included in the research studies (Reimann, 2016).

Most importantly, learning analytics is considered applied research. As such, the research intentions necessitate interdisciplinary combinations linked to both understanding and optimising the learning process. From a practical and administrative perspective, the optimisation of learning (that is, providing means for ensuring the effectiveness and efficacy of the process of learning) in part reflects the challenges education institutions now face in demonstrating quality and accountability amidst growing economic pressures (Colvin et al., 2016; Ferguson, 2012).

3. The genealogy of learning analytics

The concept of learning analytics can be traced back to the work of Pressey (1927) who developed the first automated teaching machine in the 1920s. The work of Pressey (1927) can be argued as the start of intelligent tutoring systems (ITS), one of the key areas upon from which learning analytics draws. Similarly, another critical influence has been cognitive science, which originated in the work of Miller (1956) and new advances in computer science and artificial intelligence. In 1956, the first adaptive teaching system known as the Self-Adaptive Keyboard Instructor (SAKI), was developed for teaching keyboard skills (Pask, McKinnon-Wood, & Pask, 1961). SAKI optimised learning rates by aligning the difficulty of the tasks with a learner's performance. Although by today's measures these efforts were very basic, they did serve to demonstrate how student learning can be supported through the use of technologies at scale.

An important finding that profoundly shaped the development of modern educational technology and, subsequently learning analytics, is the growing realisation of the benefits of personalised instruction. The seminal "two-sigma" study by Bloom (1984) showed that students in personalised learning condition perform one standard deviation better than students in mastery-teaching condition (the first sigma), which in turn perform one standard deviation better than students in traditional classroom-based learning condition (the second sigma)[2]. These, and similar findings, coupled with massive technological advances of the day, resulted in the significant progress within the ITS field and the field of computer-assisted instruction (CAI). Although such systems were seen to be highly advanced and innovative for their time, their specialised nature–and hence high development and production costs–presented a challenge in extending these sytems into broader adoption.

The growth of online and distance education further contributed to the development of learning analytics (Joksimović, Kovanović, Skrypnyk, et al., 2015). Starting with the use of the postal services in the late 19th Century, distance education has always been reliant on technology to reduce the barriers to effective learning and teaching (Kovanović et al., 2015). Distance education has experimented with various technologies including radio, television, video, CD, DVD, and now more commonly, the Internet (Anderson & Dron, 2010). The aim of using these technologies has been to reduce the time between students accessing content, or interacting with their teacher and other students (Joksimović, Gašević, Loughin, Kovanović, & Hatala, 2015; Moore, 1989). Moore (1993) calls the lag in accessing learning resources "transactional distance".

A key milestone in distance education history was the development of two-way communication technologies in the 1980s. Such technologies enabled the shift towards social-constructivist learning, placing a greater focus on facilitating quality interactions between students and instructors rather than the simple transmission of information. The development of World-Wide Web in the 1990s gave birth to Web-based distance learning systems which in turn evolved into modern-day online learning (Harasim, 2000; Joksimović, Kovanović, Skrypnyk, et al., 2015). The expansion of Internet and Web-based technologies ultimately resulted in the development of Massive Open Online Courses (MOOCs), a particular form of online learning in which thousands, and even hundreds of thousands, of students engage in distributed, online learning.

The expansion of Internet during the 1990s and 2000s led to web-based distance learning technologies known as Learning Management Systems (LMS), becoming increasingly used to support traditional, brick-and-mortar classroom-based learning (Harasim, 2000; Joksimović, Kovanović, Skrypnyk, et al., 2015). The broader adoption of such technologies beyond distance education provided new forms of student engagement, with teachers increasingly incorporating online activities and assessments into their face-to-face classroom teaching. This gave birth to new, blended, modes of learning characteristic of the present-day learning environment (Skrypnyk, Joksimović, Kovanović, Dawson, et al., 2015).

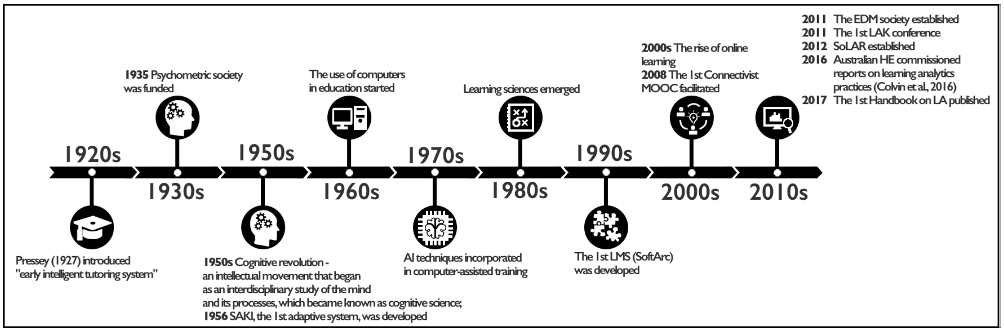

Figure 1. The genealogy of Learning Analytics.

While similar to intelligent tutoring systems (ITS), the LMS and similar technologies, are more open and flexible than previous systems, allowing for a greater range of diversity in teaching approaches, contexts and disciplines (Coates, James, & Baldwin, 2005; Weaver, Spratt, & Nair, 2008). Because staff required only a minimal set of technical skills to create an online course the use of an LMS reduced development and production costs. As such, the "ease of use" of LMS-based technologies has allowed for the rapid expansion into all facets of education. In contrast, the high technical skills associated with ITS and the closed and context-specific nature of the technology militated against wider sector uptake.

4. From early predictions to multimodal learning analytics

The use of learning technologies in distance and face-to-face teaching resulted in the collection of vast amounts of learning-related data. As noted in the previous section, over time the growing adoption and sophistication of educational technologies in learning and teaching have provided for a parallel pursuit in the use and analysis of student data. Initially, the data analytics related to aspects of web usage statistics to evaluate uptake or impact of a tool as well as basic business intelligence reports from student admissions and enrolment numbers (Clow, 2013; Dawson, Heathcote, & Poole, 2010; Dawson, McWilliam, & Tan, 2008). The broad-scale uptake of LMSs provided tremendous new opportunities to bring together data analytics, learning design and technology to rethink and develop the models for adaptive and personalised learning (Gašević, Dawson, & Siemens, 2015; Siemens, 2013). With respect to learning analytics, the development of personalised learning first stems from the ability to predict learning success and identify learners at risk from the analyse of trace data stored by technologies such as LMSs (Gašević et al., 2015). Such efforts have positioned learning analytics as a methodology or tool to address concerns surrounding student retention and in turn, provide a substantial economic benefit to the student and his or her institution. However, the range and abundance of data did surface many new technical and social challenges in the teaching and learning domain.

Predictive analytics: Supporting student learning by predicting future

Learning analytics is concerned with both *understanding* and *optimising* learning (Macfadyen and Dawson 2010). It is, therefore, of little surprise that much of the early research predominantly focused on establishing predictive models of student retention and academic performance (Gašević et al., 2015; Siemens, 2013), particularly as identifying or predicting students at risk of academic failure early in their academic candidature has the economic incentives of retaining students.

One of the most highly cited examples of an early learning analytics tool is designed to aid instructors provide feedback to students based on their predicted success (Arnold and Pistilli, 2012). The tool called *Course Signals* consisted of a predictive model for detecting students' at-risk of course failure, and a dashboard which uses a traffic light analogy to visualise individual students' risk of failure (i.e., green–no risk, yellow–moderate risk, red–high risk). The predictive model underpinning the *Course Signals* software is based on a wide range of variables including, LMS engagement activity, demographics, and past academic performance. The use of the predictive model and associated visualisations acts as a scalable solution for providing early and timely personalised feedback. However, as with many learning analytics tools and models, the reality of implementation does not always reflect the initial potential nor intention. This is well noted by Tanes and colleagues (2011), who showed that despite a teachers' intention to provide summative feedback using *Course Signals*, there was a tendency to frame such feedback in simplistic terms resulting in a lack of student action on the provided feedback. In contrast, teachers who manually included more actionable insights as part of the *Course Signals* feedback were more likely to improve students' learning outcomes. This highlights that the outcomes of learning analytics manifest within social systems and as such, the process of technical development has to take into account the challenges of adoption and application in real-world settings.

Social learning analytics: understanding student interactions through social network analysis.

In addition to the commonly used data sources about students and their individual learning strategies, data about students' social interactions with their peers and teachers have also attracted significant attention of learning analytics researchers (Dawson, 2008; Ferguson & Shum, 2012). Social network analysis (SNA) quickly emerged as one of the cornerstones of the learning analytics research (Dawson et al., 2014). SNA has long been a prominent method in educational research (Dawson, Bakharia, Heathcote, & others, 2010; Haythornthwaite, 1996). However, within learning analytics, the crucial difference from other educational research is the opportunity to automatically extract large-scale networks from learners' interactions across various environments, such as LMSs and different social media platforms (e.g., Twitter, Facebook). SNA work in learning analytics has involved the extraction of peer interactions evolving from online forums use to provide indicators of student sense of community (Dawson, 2008), creative capacity (Dawson, 2010), understand the association between learners' social centrality and learning outcome (Dowell et al., 2015; Joksimović, Dowell, et al., 2016) or visualize and examine regularities in interactions emerging from social learning activities that students and teachers engage with (Schreurs, Teplovs, Ferguson, de Laat, & Buckingham Shum, 2013; Skrypnyk, Joksimović, Kovanović, Gašević, & Dawson, 2015), to name a few.

While these studies provided for new avenues of investigation it was recognised that the simplistic accounting of interaction between peers does not necessarily equate to a focus on learning. Moreover, findings from the studies exploring the association between social centrality and academic outcome were often inconsistent or even contradictory. Here, Joksimović, Manataki, et al. (2016), Poquet and Dawson (2016), or Zhu and colleagues (2016), among others, began to examine not only the presence of interactions in a network but the basis for the developed relationships. As the field evolved, more sophisticated methods for statistical modelling of network dynamics and formation were used to further examine the nature of social mechanisms that drive the formation of social networks among students and the factors that influence the formation of student social networks across various formal and informal learning settings.

Discourse analytics: Understanding student communications

In addition to the analysis of structured and straightforward educational log data, the expansion of educational technologies produced vast amounts of unstructured, textual data about student learning. The field of discourse analytics (DA) (C. Rosé, 2017) is a type of learning analytics that focuses on using textual discourse data for supporting student learning. While DA techniques can be used for analysis of all kinds of textual data (e.g., student essays, open-ended responses), it is primarily used for analysis of student online communications such as transcripts of student online discussions, chat rooms, and communications from various kinds of social media (e.g., *Twitter*, *Facebook*, blogs). For example, Joksimović et al. (2015) used discourse analysis to examine the difference in student asynchronous communication across Facebook, Twitter and blogs in a large connectivist MOOC (cMOOC). Discourse analytics have also been extensively used to examine students' synchronous communication, such as the use of online chat platforms and to provide support via automated chat agents (Ferschke, Yang, Tomar, & Rosé, 2015; C. P. Rosé & Ferschke, 2016). In both scenarios, analysis of student communication transcripts and linguistically modelling student dialogue provides ways of capturing social aspects of student learning.

The important characteristic of discourse analytics is extensive use of natural language processing techniques (Kao & Poteet, 2007) for extracting quantitative measures from written text. The extracted metrics are then used for further processing by different machine learning algorithms. For example, many discourse analytic systems make use of N-grams, which are simple metrics that capture how many times textual chunks of N-words appear in a given text. For instance, Kovanović et al. (2014) used *Stanford CoreNLP toolkit* (Manning et al., 2014) to extract unigrams, bigrams and trigrams from student discussion posts as metrics for capturing the development of student critical thinking in online discourse.

The same method has been used for other learning analytics problems, such as understanding student reflective writing (Kovanović et al., 2018; Ullmann, 2017), online dialogue (Ezen-Can, Grafsgaard, Lester, & Boyer, 2015; Rebecca Ferguson, Wei, He, & Buckingham Shum, 2013) and identification of content-related online discussions (Cui & Wise, 2015). Building upon simple metrics of written text, different tools, such as *Coh-Metrix* (Graesser, McNamara, & Kulikowich, 2011) and *Linguistic Inquiry and Word Count* (LIWC) (Tausczik & Pennebaker, 2010) are used to define metrics of different discourse and phycological processes that are more generalisable than simple N-grams. These tools have been extensively used within discourse analytics for a wide range of problems, including writing quality and feedback (Crossley, Roscoe, & McNamara, 2014; McNamara, Graesser, McCarthy, & Cai, 2014; McNamara et al., 2012; Snow, Allen, Jacovina, Perret, & McNamara, 2015), online learning in MOOCs (Dowell et al., 2015; Fincham et al., 2019), online discussion engagement (Kovanović et al., 2016; Yoo & Kim, 2013), and self-reflection (Kovanović et al., 2018; Ullmann, 2017). In contrast to very context-specific, and therefore less generalisable N-grams measures, metrics extracted by *Coh-Metrix* and *LIWC* are more generalisable to other contexts, leading to analytical models which achieve better performance on new, previously unseen data.

Learning design: A missing piece in learning analytics

A limitation of early learning analytics predictive models was the lack of consideration for specific learning contexts. The developed models were scaled across teaching contexts and therefore frequently failed to take into account the specifics of particular course designs (Gašević, Dawson, Rogers, & Gašević, 2016). However, as research on developing predictive models increased the focus began to shift towards models that could account for course learning designs (Lockyer, Heathcote, & Dawson, 2013; Rienties, Toetenel, & Bryan, 2015).

The design of learning activities has a substantial effect on the translation of predictive models across different courses (Gasevic et al., 2016; Lockyer et al., 2015). In particular, the types of learning interactions (i.e., student-student, student-teacher, student-content, or student-system) that are designed into the course have critical importance on student learning outcomes (Joksimović, Gašević, et al., 2015). Joksimović, Gašević, and colleagues (2015) demonstrated that the establishment of a predictive model drawn from aggregated data across multiple courses can over or underestimate certain predictive factors when considered at the individual course level. In short, the learning design of a course needs to be factored into the development of any predictive model. The study by Joksimović, Gašević, et al. (2015) provides two key findings. First, the need for learning analytics to transition from generic to specific models. Second, the complexity associated with deploying such predictive models for application by teaching staff is an important area for further investigation.

Recently, the development of advanced sensing and machine learning technologies has given rise to new forms of learning analytics, known as multimodal learning analytics. This new form of learning analytics is receiving increasing attention as an approach able to provide more specific learning models to account for alternate learning designs and teaching practices. The multimodal learning analytics utilises rich data sources to better describe leaning in various settings–ranging from face-to-face to online educational contexts (D'Mello, Dieterle, & Duckworth, 2017; Martinez-Maldonado et al., 2016; Spikol et al., 2017). Multimodal learning analytics tends to capture complementary sources of learning related data, providing the basis for a robust understanding of learning processes (Ochoa, 2017). Such approaches tend to go beyond more traditional trace and survey data to incorporate various sensor data streams that capture gestures, gaze, or speech (Azevedo, 2015; Ochoa, 2017). The focus on the use of sensors in capturing various aspects of students' engagement and learning processes indicates that such studies are usually conducted in the laboratory settings, given the very limited scalability of those devices and their applicability in the context of the traditional classroom. However, the current trends in multimodal learning analytics (e.g., *CrossMML*[3]) research reflects the tendency to transition towards the real-world studies. Nevertheless, challenges remain with respect to data synchronisation, having data streams coming from various platforms and the different set of devices (Shankar, Prieto, Rodriguez-Triana, & Ruiz-Calleja, 2018). In this regard, considerable efforts have been devoted to the development of software architectures that would allow for seamless integration of multiple data streams (Shankar et al., 2018).

5. Moving learning analytics forward

Current trends in learning analytics research have tended to focus less on the development of technologies and more on the theory and principles of teaching and learning. The following outlines four promising areas of investigation.

Learning analytics for supporting student learning

To date research and development in learning analytics feedback and dashboards has been focused more towards teaching staff in lieu of personalised student facing analytics (Jovanović et al., 2017). With large-scale data at hand, it became apparent that identifying patterns in underlying data and predicting potential outcomes was not enough (Duval, 2011). It is also critical to identify personalised approaches for presenting learning data, in a way that builds upon the students existing educational knowledge and practices and does not produce information overload (Chatti, Dyckhoff, Schroeder, & Thüs, 2012). In a recent systematic review of research on learning analytics dashboards, Matcha et al (2019) noted that such works are not driven by existing educational theories and simply provide a presentation of readily available data. Matcha and colleagues (2019) further note that current learning analytics dashboards also fail to support the development of metacognitive skills, do not offer information about effective learning tactics and strategies, and cause significant concerns with respect to their evaluation. Thus, there are growing calls for the grounding of learning analytics dashboards in the "literature on learning processes, effective study methods, and feedback" (Matcha et al., 2019, p. 17).

Grounding learning analytics in educational theory

A common criticism of learning analytics is the lack of theoretical underpinning of its research. For example, development of predictive models of student success and retention described in the previous section relies on simple proxies of learning (Bergner, 2017; Dawson, Mirriahi, & Gasevic, 2015; Wise & Shaffer, 2015). Student trace data are essentially recorded counts within a specific technology. To understand what constitutes a meaningful measure of learning requires integrating relevant theory to the associated analytics (Bergner, 2017; Knight & Buckingham Shum, 2017). Bergner (2017) makes a critical distinction between predictive and explanatory models, arguing for the importance of understanding the difference between the algorithmic modelling culture and theory-driven view. Bergner (2017) argued that,

> while an explanatory model can be used to make predictions—and an error-free explanatory model would make perfect predictions—a predictive model is not necessarily explanatory" (p.42).

Predictive modelling aims to reduce bias and variance and therefore, often sacrifices “theoretical accuracy for improved empirical precision" (Shmueli, 2010, p. 293). However, to obtain actionable insights that would allow for advancing the process of learning, "it is explanatory power that plays this role" (Bergner, 2017, p. 42). For example, the application of neural-networks to predict academic performance may do little to explain why students are failing. Alignment of models to learning theory can provide for deeper more practical insights (explanatory power) for teachers to act on.

The importance of theory in learning analytics also stems from the notion of validity in educational measurement (Bergner, 2017; Joksimović et al., 2018). Validity is viewed as the degree to which theory and evidence support the interpretation of the measurement. According to Kane (2006), performance assessment should not be restricted to test items or test-like tasks and should instead include a wide variety of tasks, performed in different contexts and situations. For instance, accurately assessing student performance in MOOCs requires taking into consideration how evaluation metrics were defined in a particular learning environment (Kane, 2006; Moss, Girard, & Haniford, 2006). In that sense, Joksimović and colleagues

(2018) provide a comprehensive alignment between learning analytics and theories of learning. Specifically, in addition to positioning MOOCs as informal digital education used to facilitate learning at scale, Joksimović and colleagues (2018) provide a re-operationalisation of commonly used metrics about specific educational variables, learning context, learning processes, and learning outcomes. From a teaching and learning perspective, the findings from this study provides an understanding of how evaluation metrics have been defined in the context of MOOCs, enabling teachers to make actionable interpretations of student performance in the context of their specific learning setting (Kane, 2006; Moss et al., 2006).

Learning analytics for feedback provision and instructional interventions

With respect to educational assessment, the focus of learning analytics is primarily on formative assessment for learning (i.e., the assessment focused on improving student learning) and assessment as learning (i.e., assessment as a specific learning activity), rather than typical summative assessment of learning (i.e., assessment as a measurement of student's knowledge) (Knight, Buckingham Shum, & Littleton, 2013). This primarily stems from learning analytics methods and techniques providing timely, actionable, and personalised insights to students and teachers (Jovanović, Gašević, Dawson, Pardo, & Mirriahi, 2017). In this regard, significant work has been done on moving beyond grades and identifying students at risk (Hlosta, Zdrahal, & Zendulka, 2017; Robinson, Yeomans, Reich, Hulleman, & Gehlbach, 2016), to providing means to measure critical thinking (Kovanović, 2017), creativity (Peng, Cherng, & Chen, 2013), collaboration, and other higher-level processes (Marbouti & Wise, 2016; Wen, Yang, & Rose, 2014).

Learning analytics for understanding student emotions

Recent literature shows that emotion is one of the fundamental elements impacting on learning in online settings (Kozan & Caskurlu, 2018; Stenbom, Cleveland-Innes, & Hrastinski, 2016). According to D'Mello (2017) every learning activity is underpinned with certain emotional responses, being positive (e.g., joy, pride, satisfaction) or negative (e.g., anger, frustration, anxiety) Several attempts were made to extend the most commonly used approach of analysing trace data to understand learning processes to extract affective dimensions from students' interactions with technology (D'Mello et al., 2017). A limitation in this research is that studies of the association between trace data metrics and emotions are usually conducted in a laboratory setting, where affective states (such as anger, anxiety, or boredom) are captured using various judgment protocols (Porayska-Pomsta, Mavrikis, D'Mello, Conati, & Baker, 2013) or self-reports (D'Mello, 2012). Promising new directions have come from the overlap between research on affect and emotions in learning analytics and the research in multimodal learning analytics where attempts were made to detect affect from body signals, using various protocols for classroom observations or coding recorded interactions (D'Mello, 2017).

6. Putting learning analytics into practice

With the rapid growth of interest in learning analytics the field continues to mature in all aspects of its analytical methods and techniques, application into practice, and theoretical contributions (Dawson, Drachsler, & Rose, 2016). Nevertheless, there remains a significant a gap in the research concerning learning analytics adoption in higher education institutions (Colvin et al., 2016; Tsai & Gasevic, 2017). According to recent reports, the majority of the institutions are aware of the benefits provided by the analysis of large-scale data about student learning (Colvin et al., 2016; Tsai & Gasevic, 2017). Yet, the use of learning analytics remains mostly limited to the basic reporting about student engagement. This shows that most institutions are in the early stages of adoption and are still to realise the potential learning analytics can bring to an organisation (Haythornthwaite, 2017; Tsai & Gasevic, 2017).

From the early work of Goldstein and Katz (2005), Bichsel (2012) and Norris and Baer (2012) there has been an ongoing development of adoption models to aid the uptake of analytics in university settings. Bichsel developed the *Maturity Index* to benchmark effective institutional adoption against relevant dimensions. Bichsel argued that for analytics adoption universities must address aspects related to the organisation's culture and processes; ability to access and report on data; long-term investment in staff expertise and infrastructure; as well as overarching models for governance. Oster, Lonn, Pistilli, and Brown (2016) later revised this work to develop the *Learning Analytics Readiness Instrument* (LARI). The dimensions comprising the LARI closely reflect Bichsel's previous work. However, differences in the two models lie in the framing of the instruments for scaling learning analytics. The LARI instrument was designed to aid the preparation of organisational learning analytics deployment in lieu of evaluating or benchmarking the progression of analytics adoption. While these models have actively contributed to the discussion surrounding the primary dimensions impacting organisational adoption of learning analytics they are limited by their oversimplification of the inter-relationships between the dimensions.

Greller and Drachsler's (2012) research on learning analytics adoption models presents an alternate framing that attempts to capture the recursive nature of the critical factors influencing organisational adoption. Herein the authors noted the importance of leadership, uniting multiple stakeholders, and the development of privacy and ethics legislation—all framed within a broader strategic framework. While the Greller and Drachsler's (2012) model further extends our understanding of the complexity of the intersecting dimensions it does little to articulate how organisations can start to transition learning analytics from the classroom to the whole organisation. Ferguson and colleagues (2014) argued for the use of *RAPID Outcomes Mapping Approach* (ROMA). A noted feature of the ROMA model is the importance of identifying the key agents involved in large-scale adoption processes. As such the ROMA model begins to take on more of a systems perspective to illustrate that the adoption of complex entities—such as learning analytics—is non-linear and frequently presents numerous unpredicted outcomes.

Colvin et al's (2016) study of learning analytics adoption in Australian higher education demonstrated empirically the conventional approaches for deployment from a dynamic system view (see Figure 2).

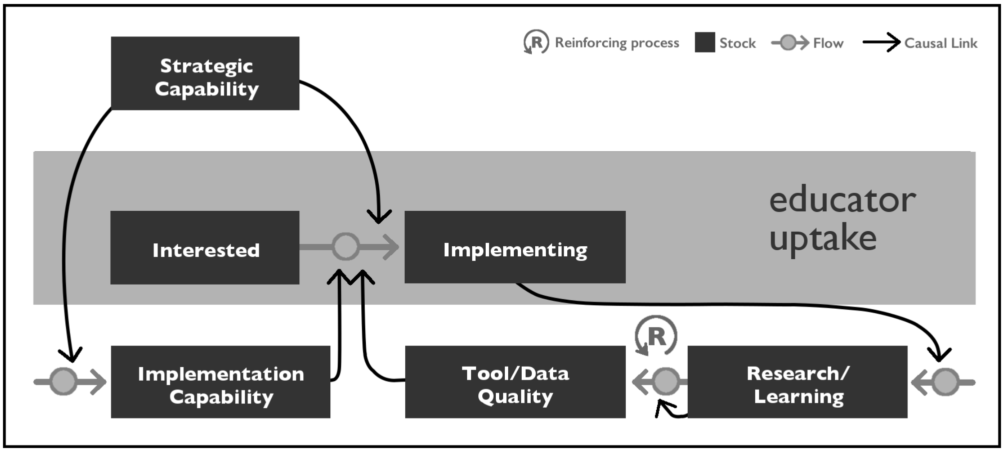

Figure 2. How Australian universities have adopted learning analytics (from Colvin et al., 2016).

Colvin and colleagues (2016) identified that institutions implemented learning analytics to either resolve an identified challenge (e.g. student retention) or a process to improve learning and teaching practice through promoted small-scale innovations. While such solution-focused instantiations provided a foundation for learning analytics use, the approach lacked sufficient adaptivity and responsiveness to engage all stakeholders in order to address an issue as multifaceted and complex as retention. Conversely, the second trajectory identified by Colvin and colleagues (2016) covered the core agents and components of the system but further leadership and identification of critical strategic outcomes are required to transition from small to large scale. Notable in this process is the recognition that learning analytics is multi-disciplinary and touches on all facets of education—from technical infrastructure, teaching quality, student learning experience, assessment practices and workload models to name but a few.

To address the complexity of scaling learning analytics Dawson and colleagues (2019, 2018) argued for the inclusion of new forms of leadership models in education to stimulate and promulgate systemic change. This remains an under-examined area in the learning analytics field. Further research is required to unpack not only the leadership attributes and approaches that can enable learning analytics instantiations but also to review and examine the processes that enable LA to move from small to systemic scales.

7. Summary

Since the emergence of first LMSs we have witnessed a proliferation of various platforms used to support online (and blended) course delivery (Siemens, Gašević, & Dawson, 2015). The fact that these platforms were primarily designed to *support course delivery*, and not necessarily empowering *design for learning* (Carvalho & Goodyear, 2014), is a major limitation in collecting data that would better reflect learning, either as a process or as an outcome. The existing research argues for the importance of learning analytics to move beyond "mere clicks" (Gašević et al., 2015) in order to provide theoretically sound interpretations of students' interactions with the underlying learning environments. The goal of this paper has been to present a brief narrative of the history of learning analytics that outlines the progression of the field and the significant contributions that emerge when disparate disciplines come together.

As detailed above, research in learning analytics has rapidly progressed from studies developing predictive models of student retention to more acute challenges linked to affect, self-regulation and feedback processes. Dawson and colleagues (2018) highlighted the increasing tensions arising in learning analytics between the growing sophistication of research derived from small-scale studies and the ability to translate these findings at scale.

Despite the popularity of learning analytics, increasing availability of data and learning analytics tools as well as the ongoing noted importance of learning analytics in education there remains significant barriers and challenges in organisational adoption. As Dawson and colleagues (2018) noted: "while LA research is rapidly, yet independently, progressing, education institutions remain mired in a quagmire of technical, social and cultural melees" (p. 236). In short, the field suffers from significant translational and contextual problems.

Learning analytics is applied research, and as such, there is much potential in the theory it generates. The value of learning analytics lies in the ability to provide more timely and personalised feedback and learning pathways at scale. In short, learning analytics increases the quality and quantity of feedback loops in the education system for teachers, learners and administrators. For education to respond to the complex sets of drivers in modern society (e.g. artificial intelligence, workforce reskilling, government funding, policy changes, industry partnerships, diverse and changing cohorts, education costs and lifelong education requirements etc.) there is a dire need for analytics to be connected with applied and practical feedback loops. However, this potential will only be realised through the convergence of technical and social systems. This requires extending current learning analytics research from technical approaches (e.g. tool development and assessment) to investigations of the social system that develop a better understanding of how learning analytics are adopted and applied in complex education systems.

To realise its full potential learning analytics has to be understood as a continual process of incremental improvement and evolution rather than a one-off effort (Rubenstein-Montano et al., 2001). In that sense, similar to the field of knowledge management a few decades ago, we need to position learning analytics in a broader context of systems thinking (Dawson et al., 2019). By learning from the work of Rubenstein-Montano and colleagues (2001), such a conceptualisation would have several important implications for understanding and adopting learning analytics. Firstly, in order to allow for successful adoption, institutional strategies and goals must be underpinned by learning analytics principles. Secondly, to address the needs of various stakeholders, we need to plan (e.g., when designing a course) before undertaking specific learning analytics activities. Finally, (and similar to knowledge management) learning analytics can be observed as an "evolutionary, iterative process directed by feedback loops and learning" (Rubenstein-Montano et al., 2001, p. 13).

Although interactions recorded within the LMS are an invaluable proxy for understanding learning, metrics extracted from trace data do not necessarily align with contemporary learning theories. Any lack of alignment will make it even more challenging to design learning tasks that would yield learning activities[4] which further inform actionable insights for impacting teaching and learning.

8. Notes

1. Information on the Society for Learning Analytics Research (SoLAR) and its publications is available online at https://solaresearch.org/about/

2. In this context, sigma refers to standard deviation, which is in statistics commonly represented by the small Greek letter sigma σ.

3. Information on *Multimodal Learning Analytics Across (Physical and Digital) Spaces* (CrossMMLA) workshop is online at http://crossmmla.org/

4. Activity here is defined according to Goodyear and Carvalho (2014) as being emergent from the designed learning tasks.

9. References

Anderson, T., & Dron, J. (2010). Three generations of distance education pedagogy. *The International Review of Research in Open and Distance Learning*, *12*(3), 80–97.

Arnold, K. E. ., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *ACM International Conference Proceeding Series* (pp. 267–270). https://doi.org/10.1145/2330601.2330666

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84–94. https://doi.org/10.1080/00461520.2015.1004069

Baker, R. S. & Inventado, P. S. (2014). Educational Data Mining and Learning Analytics. In A. J. Larusson & B. White (Eds.), *Learning Analytics: From research to practice* (pp. 61–75). New York, NY: Springer New York. Retrieved from http://dx.doi.org/10.1007/978-1-4614-3305-7_4

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–17.

Bergner, Y. (2017). Measurement and its uses in learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The Handbook of Learning Analytics* (pp. 34–48). Alberta, Canada: Society for Learning Analytics Research (SoLAR). Retrieved from http://solaresearch.org/hla-17/hla17-chapter1

Bichsel, J. (2012). *Research computing: The enabling role of information technology.* Louisville, CO: EDUCAUSE Center for Applied Research.

Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher, 13*(6), 4–16. https://doi.org/10.2307/1175554

Carbonell, J. R. (1970). AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems, 11*(4), 190–202. https://doi.org/10.1109/TMMS.1970.299942

Carvalho, L., & Goodyear, P. (2014). *The architecture of productive learning networks*. New York, NY:Routledge.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5–6), 318–331. https://doi.org/10.1504/IJTEL.2012.051815

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education, 18*(6), 683–695. https://doi.org/10.1080/13562517.2013.827653

Coates, H., James, R., & Baldwin, G. (2005). A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary Education and Management, 11*(1), 19–36. https://doi.org/10.1007/s11233-004-3567-9

Colvin, C., Rogers, T., Wade, A., Dawson, S., Gasevic, D., Buckingham Shum, S., … Fisher, J. (2016). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*. Sydney: Office for Learning and Teaching. Retrieved from https://opus.lib.uts.edu.au/handle/10453/117173

Crossley, S., Roscoe, R., & McNamara, D. S. (2014). What is successful writing? An Investigation into the multiple ways writers can write successful essays. *Written Communication, 31*(2), 184–214. https://doi.org/10.1177/0741088314526354

Cui, Y., & Wise, A. F. (2015). Identifying content-related threads in MOOC discussion forums. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale* (pp. 299–303). New York, NY, USA: ACM. https://doi.org/10.1145/2724660.2728679

Dawson, S. (2008). A study of the relationship between student social networks and sense of community. *Educational Technology & Society, 11*(3), 224–238.

Dawson, S., Bakharia, A., Heathcote, E., & others. (2010). SNAPP: Realising the affordances of real-time SNA within networked learning environments. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the 7th International Conference on Networked Learning* (pp.125-133). Aalborg, Denmark: Aalborg University.

Dawson, S., Drachsler, H., & Rose, C. P. (Eds.) (2016). *LAK '16: Proceedings of the sixth international conference on learning analytics & knowledge*. New York, NY, USA: ACM.

Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (pp. 231–240). New York, NY, USA: ACM. https://doi.org/10.1145/2567574.2567585

Dawson, S., Heathcote, L., & Poole, G. (2010). Harnessing ICT potential: The adoption and analysis of ICT systems for enhancing the student learning experience. *International Journal of Educational Management*, *24*(2), 116–128. https://doi.org/10.1108/09513541011020936

Dawson, S., Joksimović, S., Poquet, O., & Siemens, G. (2019). Increasing the impact of learning analytics. In *Proceedings of the Ninth International Conference on Learning Analytics & Knowledge (LAK'19)*. New York, NY, USA: ACM. https://doi.org/10.1145/3303772.3303784

Dawson, S., McWilliam, E., & Tan, J. P.-L. (2008). Teaching smarter: How mining ICT data can inform and improve learning and teaching practice. In *Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 221–230). Melbourne, Australia: Deakin University. Retrieved from https://ro.uow.edu.au/medpapers/141

Dawson, S., Mirriahi, N., & Gasevic, D. (2015). Importance of Theory in Learning Analytics in Formal and Workplace Settings. *Journal of Learning Analytics*, *2*(2), 1–4.

Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gasevic, D. (2018). Rethinking Learning Analytics Adoption Through Complexity Leadership Theory. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 236–244). New York, NY, USA: ACM. https://doi.org/10.1145/3170358.3170375

D'Mello, S. (2012). Monitoring Affective Trajectories During Complex Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2325–2328). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-1428-6_849

D'Mello, S. (2017). Emotional Learning Analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The Handbook of Learning Analytics* (1st ed., pp. 115–127). Alberta, Canada: Society for Learning Analytics Research (SoLAR). Retrieved from http://solaresearch.org/hla-17/hla17-chapter1

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, Analytic, Automated (AAA) Measurement of Engagement During Learning. *Educational Psychologist*, *52*(2), 104–123. https://doi.org/10.1080/00461520.2017.1281747

Dowell, N., Skrypnyk, O., Joksimović, S., Graesser, A. C., Dawson, S., Gašević, D., … Kovanović, V. (2015). Modeling Learners' Social Centrality and Performance through Language and Discourse. Presented at the In Proceedings of the 8th International Conference on Educational Data Mining, Madrid, Spain.

Duval, E. (2011). Attention Please!: Learning Analytics for Visualization and Recommendation. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 9–17). New York, NY, USA: ACM. https://doi.org/10.1145/2090116.2090118

Ezen-Can, A., Grafsgaard, J. F., Lester, J. C., & Boyer, K. E. (2015). Classifying Student Dialogue Acts with Multimodal Learning Analytics. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp. 280–289). New York, NY, USA: ACM. https://doi.org/10.1145/2723576.2723588

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5–6), 304–317. https://doi.org/10.1504/IJTEL.2012.051816

Ferguson, R.a, Clow, D. ., Macfadyen, L. ., Essa, A. ., Dawson, S. ., & Alexander, S. . (2014). Setting learning analytics in context: Overcoming the barriers to large-scale adoption. In *ACM International Conference Proceeding Series* (pp. 251–253). https://doi.org/10.1145/2567574.2567592

Ferguson, R.a, & Shum, S. B. . (2012). Social learning analytics: Five approaches. In *ACM International Conference Proceeding Series* (pp. 23–33). https://doi.org/10.1145/2330601.2330616

Ferguson, Rebecca, Wei, Z., He, Y., & Buckingham Shum, S. (2013). An Evaluation of Learning Analytics to Identify Exploratory Dialogue in Online Discussions. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 85–93). New York, NY, USA: ACM. https://doi.org/10.1145/2460296.2460313

Ferschke, O., Yang, D., Tomar, G., & Rosé, C. P. (2015). Positive Impact of Collaborative Chat Participation in an edX MOOC. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in Education* (pp. 115–124). Springer International Publishing.

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., Staalduinen, J.-P. van, & Gašević, D. (2019). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In *Proceedings of the 9th International Conference on Learning Analytics and Knowledge*. New York, NY, USA: ACM.

Gašević, D., Dawson, S., Rogers, T., & Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68–84. http://dx.doi.org/10.1016/j.iheduc.2015.10.002

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learning: Research and Practice*, *3*(1), 63–78. https://doi.org/10.1080/23735082.2017.1286142

Goldstein, P. J., & Katz, R. N. (2005). *Academic analytics: The uses of management information and technology in higher education* (Vol. 8). Educause.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, *40*(5), 223–234. https://doi.org/10.3102/0013189X11413260

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology and Society*, *15*(3), 42–57.

Harasim, L. (2000). Shift happens: online education as a new paradigm in learning. *The Internet and Higher Education*, *3*(1–2), 41–61. http://dx.doi.org/10.1016/S1096-7516(00)00032-4

Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library & Information Science Research*, *18*(4), 323–342. https://doi.org/10.1016/S0740-8188(96)90003-1

Haythornthwaite, C. (2017). An Information Policy Perspective on Learning Analytics. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 253–256). New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3027389

Hlosta, M., Zdrahal, Z., & Zendulka, J. (2017). Ouroboros: Early Identification of At-risk Students Without Models Based on Legacy Data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 6–15). New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3027449

Joksimović, S., Dowell, N., Skrypnyk, O., Kovanović, V., Gašević, D., Dawson, S., & Graesser, A. C. (2016). Exploring Development of Social Capital in a cMOOC Through Language and Discourse. *The Internet and Higher Education*, *(revisions submitted)*.

Joksimović, S., Gašević, D., Loughin, T., Kovanović, V., & Hatala, M. (2015). Learning at distance: Effects of interaction traces on academic achievement. *Computers and Education*, *87*, 204–217. https://doi.org/10.1016/j.compedu.2015.07.002

Joksimović, S., Kovanović, V., Jovanović, J., Zouaq, A., Gašević, D., & Hatala, M. (2015). What do cMOOC participants talk about in social media? A topic analysis of discourse in a cMOOC. In *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge (LAK'15)* (pp. 156–165). New York, NY, USA: ACM. https://doi.org/10.1145/2723576.2723609

Joksimović, S., Kovanović, V., Skrypnyk, O., Gašević, D., Dawson, S., & Siemens, G. (2015). The History and State of Online Learning. In *Preparing for the digital university: a review of the history and current state of distance, blended, and online learning* (pp. 93–132). http://linkresearchlab.org/PreparingDigitalUniversity.pdf.

Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & de Kereki, I. F. (2016). Translating Network Position into Performance: Importance of Centrality in Different Network Configurations. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 314–323). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883928

Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., … Brooks, C. (2018). How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. *Review of Educational Research*, *88*(1), 43–86. https://doi.org/10.3102/0034654317740335

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*, 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001

Kane, M. (2006). Validation. *Educational Measurement*, *4*(2), 17–64.

Kao, A., & Poteet, S. R. (2007). *Natural language processing and text mining*. London: Springer.

Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., & Haydock, M. (2012). Analytics: The widening divide. *MIT Sloan Management Review*, *53*(2), 1.

Knight, S., & Buckingham Shum, S. (2017). Theory and Learning Analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The Handbook of Learning Analytics* (1st ed., pp. 17–22). Alberta, Canada: Society for Learning Analytics Research (SoLAR). Retrieved from http://solaresearch.org/hla-17/hla17-chapter1

Knight, S., Buckingham Shum, S., & Littleton, K. (2013). Epistemology, pedagogy, assessment and learning analytics. In *ACM International Conference Proceeding Series* (pp. 75–84). https://doi.org/10.1145/2460296.2460312

Kovanović, V. (2017). *Assessing cognitive presence using automated learning analytics methods* (Doctoral dissertation). The University of Edinburgh, Edinburgh, Scotland. Retrieved from https://www.era.lib.ed.ac.uk/handle/1842/28759

Kovanović, V., Joksimović, S., Gašević, D., & Hatala, M. (2014). Automated cognitive presence detection in online discussion transcripts. In *Proceedings of the Workshops at the LAK 2014 Conference co-located with 4th International Conference on Learning Analytics and Knowledge (LAK'14)*. Indianapolis, IN, USA. Retrieved from http://ceur-ws.org/Vol-1137/

Kovanović, V., Joksimović, S., Mirriahi, N., Blaine, E., Gašević, D., Siemens, G., & Dawson, S. (2018). Understand students' self-reflections through learning analytics. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK'18)* (pp. 389–398). New York, NY, USA: ACM. https://doi.org/10.1145/3170358.3170374

Kovanović, V., Joksimović, S., Skrypnyk, O., Gašević, D., Dawson, S., & Siemens, G. (2015). The history and state of distance education. In G. Siemens, D. Gašević, & S. Dawson (Eds.), *Preparing for the digital university: a review of the history and current state of distance, blended, and online learning* (pp. 9–54). Edmonton, AB: Athabasca University. Retrieved from http://linkresearchlab.org/PreparingDigitalUniversity.pdf

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., & Siemens, G. (2016). Towards Automated Content Analysis of Discussion Transcripts: A Cognitive Presence Case. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 15–24). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883950

Kozan, K., & Caskurlu, S. (2018). On the Nth presence for the Community of Inquiry framework. *Computers & Education*, *122*, 104–118. https://doi.org/10.1016/j.compedu.2018.03.010

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing Pedagogical Action: Aligning Learning Analytics With Learning Design. *American Behavioral Scientist*, *57*(10), 1439–1459. https://doi.org/10.1177/0002764213479367

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 55–60).

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity.

Marbouti, F., & Wise, A. F. (2016). Starburst: a new graphical interface to support purposeful attention to others' posts in online discussions. *Educational Technology Research and Development*, *64*(1), 87–113. https://doi.org/10.1007/s11423-015-9400-y

Martinez-Maldonado, R., Schneider, B., Charleer, S., Shum, S. B., Klerkx, J., & Duval, E. (2016). Interactive Surfaces and Learning Analytics: Data, Orchestration

Aspects, Pedagogical Uses and Challenges. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 124–133). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883873

Matcha, W., Uzir, N. A., Gaševic, D., & Pardo, A. (under review). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press.

McNamara, D. S., Raine, R., Roscoe, R., Crossley, S., Jackson, G. T., Dai, J., … others. (2012). The Writing-Pal: Natural language algorithms to support intelligent tutoring on writing strategies. In *Applied Natural Language Processing: Identification, Investigation and Resolution* (pp. 298–311).

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Moore, M. G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, *3*(2), 1–7. https://doi.org/10.1080/08923648909526659

Moore, M. G. (1993). Theory of transactional distance. In *In D. Keegan (Ed.), Theoretical principles of distance education* (pp. 22–38). New York: Routledge.

Moss, P. A., Girard, B. J., & Haniford, L. C. (2006). Chapter 4: Validity in Educational Assessment. *Review of Research in Education*, *30*(1), 109–162. https://doi.org/10.3102/0091732X030001109

Norris, D. M. ., & Baer, L. . (2012). Panel proposal: Building organizational capacity for analytics. In *ACM International Conference Proceeding Series* (pp. 18–19). https://doi.org/10.1145/2330601.2330612

Ochoa, X. (2017). Multimodal Learning Analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The Handbook of Learning Analytics* (1st ed., pp. 129–141). Alberta, Canada: Society for Learning Analytics Research (SoLAR). Retrieved from http://solaresearch.org/hla-17/hla17-chapter1

Oster, M., Lonn, S., Pistilli, M. D., & Brown, M. G. (2016). The Learning Analytics Readiness Instrument. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 173–182). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883925

Pask, G., McKinnon-Wood, R., & Pask, E. (1961). *GB866279 (A)*. European Patent Office.

Peng, S.-L., Cherng, B.-L., & Chen, H.-C. (2013). The Effects of Classroom Goal Structures on the Creativity of Junior High School Students. *Educational Psychology*, *33*(5), 540–560.

Poquet, O., & Dawson, S. (2016). Untangling MOOC Learner Networks. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 208–212). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883919

Porayska-Pomsta, K., Mavrikis, M., D'Mello, S., Conati, C., & Baker, R. S. J. d. (2013). Knowledge Elicitation Methods for Affect Modelling in Education. *Int. J. Artif. Intell. Ed.*, *22*(3), 107–140. https://doi.org/10.3233/JAI-130032

Pressey, S. L. (1927). A machine for automatic teaching of drill material. *School & Society*, *25*, 549–552.

Reimann, P. (2016). Connecting learning analytics with learning research: the role of design-based research. *Learning: Research and Practice*, *2*(2), 130–142. https://doi.org/10.1080/23735082.2016.1210198

Rienties, B., Toetenel, L., & Bryan, A. (2015). "Scaling Up" Learning Design: Impact of Learning Design Activities on LMS Behavior and Performance. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 315–319). New York, NY, USA: ACM. https://doi.org/10.1145/2723576.2723600

Robinson, C., Yeomans, M., Reich, J., Hulleman, C., & Gehlbach, H. (2016). Forecasting Student Achievement in MOOCs with Natural Language Processing. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 383–387). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883932

Rosé, C. (2017). Discourse Analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *The Handbook of Learning Analytics* (1st ed., pp. 105–114). Alberta, Canada: Society for Learning Analytics Research (SoLAR). Retrieved from http://solaresearch.org/hla-17/hla17-chapter1

Rosé, C. P., & Ferschke, O. (2016). Technology Support for Discussion Based Learning: From Computer Supported Collaborative Learning to the Future of Massive Open Online Courses. *International Journal of Artificial Intelligence in Education*, *26*(2), 660–678. https://doi.org/10.1007/s40593-016-0107-y

Rubenstein-Montano, B., Liebowitz, J., Buchwalter, J., McCaw, D., Newman, B., & Rebeck, K. (2001). A systems thinking framework for knowledge management. *Decision Support Systems*, *31*(1), 5–16. https://doi.org/10.1016/S0167-9236(00)00116-0

Schreurs, B., Teplovs, C., Ferguson, R., de Laat, M., & Buckingham Shum, S. (2013). Visualizing Social Learning Ties by Type and Topic: Rationale and Concept Demonstrator. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 33–37). New York, NY, USA: ACM. https://doi.org/10.1145/2460296.2460305

Shankar, S. K., Prieto, L. P., Rodriguez-Triana, M. J., & Ruiz-Calleja, A. (2018). A Review of Multimodal Learning Analytics Architectures. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)* (Vol. 00, pp. 212–214). https://doi.org/10.1109/ICALT.2018.00057

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310.

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Siemens, G., Gašević, D., & Dawson, S. (2015). Preparing for the digital university: A review of the history and current state of distance, blended, and online learning. *Athabasca, Canada: Athabasca University.*

Siemens, G., Long, P., Gašević, D., & Conole, G. (2011). Call for Papers, 1st International Conference Learning Analytics & Knowledge (LAK 2011). Retrieved from https://tekri.athabascau.ca/analytics/call-papers

Skrypnyk, O., Joksimović, S., Kovanović, V., Dawson, S., Gašević, D., & Siemens, G. (2015). The history and state of blended learning. In G. Siemens, D. Gašević, & S. Dawson (Eds.), *Preparing for the digital university: a review of the history and* *current state of distance, blended, and online learning* (pp. 55–92). Edmonton, AB: Athabasca University. Retrieved from http://linkresearchlab.org/PreparingDigitalUniversity.pdf

Skrypnyk, O., Joksimović, S., Kovanović, V., Gašević, D., & Dawson, S. (2015). Roles of course facilitators, learners, and technology in the flow of information of a CMOOC. *International Review of Research in Open and Distance Learning*, *16*(3), 188–217.

Snow, E. L., Allen, L. K., Jacovina, M. E., Perret, C. A., & McNamara, D. S. (2015). You've Got Style: Detecting Writing Flexibility Across Time. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp. 194–202). New York, NY, USA: ACM. https://doi.org/10.1145/2723576.2723592

Spikol, D., Prieto, L. P., Rodríguez-Triana, M. J., Worsley, M., Ochoa, X., Cukurova, M., … Ringtved, U. L. (2017). Current and Future Multimodal Learning Analytics Data Challenges. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 518–519). New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3029437

Stenbom, S., Cleveland-Innes, M., & Hrastinski, S. (2016). Emotional Presence in a Relationship of Inquiry: The Case of One-to-One Online Math Coaching. *Online Learning*, *20*(1). https://doi.org/10.24059/olj.v20i1.563

Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education*, *57*(4), 2414–2422. https://doi.org/10.1016/j.compedu.2011.05.016

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Tsai, Y.-S., & Gasevic, D. (2017). Learning Analytics in Higher Education — Challenges and Policies: A Review of Eight Learning Analytics Policies. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 233–242). New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3027400

Ullmann, T. D. (2017). Reflective Writing Analytics: Empirically Determined Keywords of Written Reflection. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 163–167). New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3027394

Weaver, D., Spratt, C., & Nair, C. (2008). Academic and student use of a learning management system: Implications for quality. *Australasian Journal of Educational Technology*, *24*(1). https://doi.org/10.14742/ajet.1228

Wen, M., Yang, D., & Rose, C. (2014). Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014* (pp. 525–534). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84909951147&partnerID=40&md5=80f121bcfc587505feaee3a3d6675c59

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.

Yoo, J., & Kim, J. (2013). Can Online Discussion Participation Predict Group Project Performance? Investigating the Roles of Linguistic Features and

Participation Patterns. *International Journal of Artificial Intelligence in Education*, *24*(1), 8–32. https://doi.org/10.1007/s40593-013-0010-8

Zhu, M., Bergner, Y., Zhang, Y., Baker, R., Wang, Y., & Paquette, L. (2016). Longitudinal Engagement, Performance, and Social Connectivity: A MOOC Case Study Using Exponential Random Graph Models. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 223–230). New York, NY, USA: ACM. https://doi.org/10.1145/2883851.2883934