

Four Paradigms in Learning Analytics: Why Paradigm Convergence Matters

Ryan S. Baker, Dragan Gasevic, Shamyia Karumbaiah

Abstract

Learning analytics has matured significantly since its early days. The field has rapidly grown in terms of the reputation of its publication venues, established a vibrant community, and has demonstrated an increasing impact on policy and practice. However, the boundaries of the field are still being explored by many researchers in a bid to determine what differentiates a contribution in learning analytics from contributions in related fields, which also center around data in education. In this paper, we propose that instead of emphasizing the examination of differences, a healthy development of the field should focus on collaboration and be informed by the developments in related fields. Specifically, the paper presents a framework for analysis how contemporary fields focused on the study of data in education influence trends in learning analytics. The framework is focused on the methodological paradigms that each of the fields is primarily based on – i.e., essentialist, entitative/reductionist, ontological/dialectical, and existentialist. The paper uses the proposed framework to analyze how learning analytics (ontological) is being methodologically influenced by recent trends in the fields of educational data mining (entitative), quantitative ethnography (existentialist), and learning at scale (essentialist). Based on the results of the analysis, this paper identifies gaps in the literature that warrant future research.

Keywords: learning analytics, artificial intelligence in education, quantitative ethnography, learning at scale, machine learning, research paradigms

1. Introduction

The year 2020 represents the 10th anniversary of the official formation of the field of learning analytics, which was initiated by the organization of the First International Conference on Learning Analytics and Knowledge (LAK) held in 2011 (Long et al., 2011). The LAK conference was established as a response to the growing opportunities for education afforded by the emergence in ‘big data’ (Siemens & Baker, 2012). The creation of LAK also aimed to bring together researchers and practitioners who had worked on topics related to big data in education but who had not had a joint community to exchange, discuss, and develop ideas (Siemens, 2014). The conference organizing committee for the first edition of LAK defined learning analytics as ‘the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs’ (Long et al., 2011). This definition is still commonly used to define learning analytics and is officially endorsed by the Society for Learning Analytics Research (SoLAR). The definition highlights the aspiration of learning analytics to use data to advance holistic understanding and enhancement of systems in education.

The field of learning analytics has matured significantly since its early days. The proceedings of the LAK conference are one of the most cited publications in the broad area of educational technologies (Google Scholar, 2020). The numbers of submissions and participants to LAK have grown steadily and LAK has become a leading conference. The learning analytics community is focused on nurturing the next generation of researchers, practitioners, and leaders through the steady LAK Doctoral Consortium series and the network of Learning Analytics Summer Institutes occurring worldwide and multiple online

lecture series. The Journal of Learning Analytics has been indexed by major journal libraries (e.g., Scopus and Emerging Sources Citation Index) and the second edition of the Handbook of Learning Analytics is underway after a successful first edition (Lang et al., 2017). SoLAR has developed a network of individual and institutional members and strong links with the educational technology industry. SoLAR has established several initiatives for field building such as several special interest groups, webinar and podcast series, newsletter, and job board.

Learning analytics has still many open challenges to resolve, some of which are very central to the identity of the field. There is still an open debate about what properties a learning analytics paper should have. Some advocate that you can not publish to LAK without an intervention. Others argue that you can not publish to LAK without data science. Addressing these perspectives is very critical for a healthy development of the field that fosters diversity of disciplinary and paradigmatic viewpoints. Specifically, it is essential to deepen our understanding of methodological paradigms that are shaping learning analytics as a field. This extends existing analysis of the progress in learning analytics that has largely been inward focused on topics and methodologies emerging in the field (Dawson et al., 2014, Dawson et al., 2019, Ferguson, 2012) or some specific sub-areas such as dashboards (Bodily & Verbert, 2018, Jivet et al., 2018, and Matcha et al., 2020), adoption (Viberg et al., 2018), and higher education (Tsai & Gašević, 2017). Others have proposed models of learning analytics that are largely focused on guiding the integration of relevant practices from other fields (Greller & Dracshler, 2012, Chatti et al., 2014, and Gašević et al., 2017, Siemens, 2013). While these reviews and models offered invaluable contributions that can shape future development of learning analytics, little research has been done to examine the links and mutual influences of learning analytics with other related fields, which are based on different methodological paradigms.

Learning analytics is not isolated from developments in sister fields, which are also dedicated to the study of data in education – in general, the use of data mining and analytics has emerged as a core topic and area of research in educational technology more broadly (Chen et al., 2020b). Several other relevant fields with somewhat different methodological perspectives and communities have been created since the start of the millennium. Educational data mining (EDM), the first of these fields, was formally initiated earlier than learning analytics. The first edition of the EDM conference was held in 2008 and the Journal of EDM was launched in 2009 following a series of workshops held at the major conferences starting in 2000. Both the conference and journal are run by the International Educational Data Mining Society (IEDMS), which connects researchers with the educational technology industry and many stakeholder groups. Two other communities with significant focus on data in education are centered around the ACM Conference on Learning @ Scale (started in 2014) and the International Conference on Quantitative Ethnography (started in 2019). However, research that offers a framework for a systematic analysis of mutual influences of these fields related to learning analytics is at best scarce. A rare exception is seen in Siemens & Baker (2012) who argued for the importance of building links between relevant fields and the Chen et al. (2020a) study that compared topics, bibliometrics, and communities between the LAK and EDM conferences. However, contemporary developments in relevant fields need to be systematically studied to deepen understanding on the emerging influences in learning analytics.

This paper presents a framework for analysis how contemporary fields focused on the study of data in education influence research trends in learning analytics. The framework is focused on methodological paradigms that each of the fields is primarily based on. Specifically, the paper uses the proposed framework to analyze how learning analytics is methodologically influenced by recent trends in the fields of EDM, quantitative ethnography, and learning at scale. Based on the results of the analysis, this paper identifies gaps in the literature that warrant future research.

2. Learning Analytics as a Practice

Learning analytics has been a field of practice as much as a field of research since its inception. The practice emphasis comes mostly due to shifts in political and economic factors (Buckingham Shum & Luckin, 2019; Ferguson, 2012; Siemens, 2013). Political and economic priorities shape the education sectors around the world. In many countries, funding models in higher education are significantly impacted by the numbers of students who successfully complete their degrees. Therefore, institutional senior managers (e.g., vice-presidents, provosts, and vice-provosts) and policy makers saw the potential of analytics to increase student retention through predictive modeling and early warning systems (Siemens, 2013). Similarly, there is increasing interest in using predictive analytics to increase graduation rates in K-12 (Bowers et al., 2012; Singh, 2018). The use of analytics to support student retention and other institutional priorities led to the formation of academic analytics (Campbell et al., 2007), a field that has strongly influenced learning analytics. This brought strong connections between the learning analytics community and the community of institutional leaders in information technology for higher education that is typically gathered around events run by EDUCAUSE.

A second key area for practice in learning analytics is the increased interest in evaluating and increasing the effectiveness of digital technologies in education. In higher education, there has been particular interest in the use of learning management systems (LMSs) (Ferguson, 2012). This has led many to explore the use of log data recorded by LMSs and create reporting systems (today known as dashboards) that were built by the use of leveraged information visualization and relatively simple data analytic techniques (Jovanovic et al., 2008; Rienties et al., 2018). In addition to offering relevant solutions for instructors, this created interest in several professional groups such as instructional/educational designers and educational/learning technologies (Weller, 2020).

Beyond learning management systems, there has been considerable interest in the use and refinement of computer-based learning environments such as adaptive learning systems, intelligent tutoring systems, and massive online open courses (Essa, 2016). One of the large factors driving this growth in K-12 has been the increasing emphasis on standardized tests, and the need for schools to demonstrate improvement on these tests in order to receive funding. This phenomena started in the USA and UK but has since spread worldwide (Lingard & Lewis, 2016). Even without a driving force like this in higher education, though, learners and instructors have flocked to these learning technologies. The user base of large platforms like ALEKS, Cognitive Tutor, Khan Academy, edX, and Coursera has grown to hundreds of thousands or millions of users. Learning analytics research has investigated how to use data to make these platforms more effective, refining knowledge assessments to improve curricular sequencing (Ritter et al., 2016), enhancing review schedules through models of human memory (Settles & Meeder, 2016), embedding sophisticated assessments into intervention strategies (Li et al., 2018), and studying what interventions work for different students and in different contexts (Sales et al., 2018). As with LMSs, there has been considerable interest in creating reports and visualizations that can inform teachers and instructors, and drive their interaction and intervention with their students (An et al., 2019; Holstein et al., 2019).

Due to this focus on practice, many learning technology vendors have been involved in events in learning analytics since the early days of the field formation, sponsoring and publishing papers in the LAK and EDM conferences since their beginnings. Indeed, today LAK has a practitioner track and EDM has an industrial track.

2.1. Fields Influencing Practice of Learning Analytics

The literature has consistently stressed the socio-technical nature of learning analytics (Dawson et al., 2019; Siemens, 2013). This is reflected in the definition of learning analytics adopted by SoLAR that learning analytics uses data about context in which learning occurs (Long et al., 2011), and the definition of educational data mining adopted by IEDMS in 2009 that the field's goal is "to better understand students, and the settings which they learn in." Several prominent authors in learning analytics note that learning analytics is used within organizational contexts where social, political, privacy, and ethical factors play at least as important role as any technological solutions (Colvin et al., 2015; Dawson et al., 2018, 2019; Gašević et al., 2017; Lynch, 2017; Macfadyen et al., 2014; Siemens, 2013). This led to significant work in learning analytics on the development of codes of practice and policy and strategy frameworks that have already influenced adoption of learning analytics in many institutions and systems (Sclater & Bailey, 2015; Tsai et al., 2018). Learning analytics as a field of practice draws significantly from fields such as organizational studies, change management, information systems, law, and ethics.

The recognition that the practice of learning analytics is situated within complex education systems has led to greater consideration of complexity science and theory (Jacobson et al., 2016). The need to position work within complex education systems is stressed as essential to maximize the practical impact of learning analytics (Dawson et al., 2019). Successful change management in learning analytics is suggested to be based on the existing understanding of complex adaptive systems (Macfadyen et al., 2014). Complexity leadership composed of operational, entrepreneurial, and enabling dimensions is seen as necessary for scalable adoption of learning analytics that promotes educational innovation (Tsai et al., 2019).

High adoption of learning analytics in practices calls for a strong focus on design and focus on human factors (Buckingham Shum et al., 2019; Gašević et al., 2017). After early enthusiasm in learning analytics to produce different tools (mostly dashboards), several studies found that the adoption of such tools is in many cases relatively low and may lead to undesirable outcomes such as decline in mastery orientation (Lonn et al., 2015) and grade point average (Chaturapruek et al., 2018) in undergraduate students if designed ineffectively (Bodily & Verbert, 2017; Corrin & de Barba, 2014). Therefore, researchers have recently proposed approaches that aim to guide the design of learning analytics systems and provide understanding of sensemaking (Holstein et al., 2019; Wise & Jung, 2019). Buckingham Shum et al. (2019) even take it a step further and argue for the strong consideration of human factors in learning analytics, a direction that clearly emphasizes that learning analytics practice is getting more influenced by design science and human computer interaction.

2.2. Models of Learning Analytics Practice

Learning analytics has several models that are conceptualized to guide practice, recognize complexity of the field, and acknowledge the above influences. Generally, there are two groups of models – models that identify critical dimensions and questions the field of practice needs to consider; and models that focus on the process of implementation of learning analytics. Chatti and colleagues (2012) propose a reference model of learning analytics. The reference model is composed of four dimensions to address these questions: what? – data, environment, and context; why? – objectives; how? – methods; and who? – stakeholders. Similar four dimensions are part of a generic framework for learning analytics suggested by Greller & Drachsler (2012), except that slightly different names and the scope for the four dimensions are used. Greller & Drachsler also add internal limitations (i.e., competencies and acceptance) and external constraints (i.e., norms and conventions) into the model.

Several process models of learning analytics have been proposed. One commonality across process models is that they conceptualize learning analytics as a cycle. Clow's (2012) learning analytics cycle starts with learners whose actions generate data. The data are then measured and analyzed, and the results finally inform interventions for learners. Other process models are more granular and recognize other activities of the learning analytics cycle. The interactive process model of learning analytics suggested by Steiner and colleagues (2014) consist of the following phases: data selection, data capturing, data aggregation, data reporting, prediction, acting upon results, and refinement. Siemens' (2013) data loop has similar phases to those of the model of Steiner et al. The main addition of Siemens' model is that it identifies relationships between the phases of the data loop with learning analytics purposes, sources of data, stakeholders, analytic techniques, and types of actions. That is, Siemens directly implies that learning analytics is an interdisciplinary field of research and practice that depends on the contributions of many types of stakeholders.

3. Four Intellectual Paradigms for Learning Analytics

3.1. The Four Paradigms

Whether learning analytics is considered as a practice or learning analytics is considered as a science, it cannot be denied that learning analytics is quite splintered for such a young community. Three primary conferences exist: Educational Data Mining (EDM; founded 2008), Learning Analytics and Knowledge (LAK; founded 2011) and ACM Learning @ Scale (L@S; founded 2014). A fourth conference, Quantitative Ethnography (QE; founded 2019) has a scope that goes beyond learning and education, but the majority of the work in its first two iterations was in learning and educational domains. A fifth conference, Artificial Intelligence and Education (AIED; founded 1989), precedes the foundation of learning analytics as an area of research or inquiry. AIED was originally conceived more broadly, in terms of all applications and uses of artificial intelligence in education, and in early years consisted heavily of work in the design of intelligent systems for education and models of students for those systems. EDM arguably formed as a splinter group off of AIED and AIED rapidly began to publish almost as much of this type of research as EDM.

Other conferences have seen the publication of considerable learning analytics research. The Intelligent Tutoring Systems (ITS) conference has published similar work to AIED throughout its history, and for many years the two conferences alternated years. The first EDM conference was in fact co-located with ITS, in Montreal in 2008. Other conferences such as the International Conference on Learning Sciences (itself an earlier offshoot from AIED) and Computer-Supported Collaborative Learning (CSCL) publish learning analytics work while having a remit that is both broader (other forms of research) and narrower (a smaller range of phenomena). Other organizations also exist, such as the (training rather than publication) Learning Analytics Summer Institute and Simon Initiative Summer School, and regional and national conferences.

Part of the first splintering – the emergence of LAK when EDM already existed – may be due to the conference founders not knowing each other. Very few of the founders of EDM and LAK had even met when the first LAK conference was organized, a fact that is curious in itself. However, we argue that the splintering is not just a coincidence due to the poorly-connected social network of scientists working in the field in that moment – instead, we argue that it is due to deeper differences in how different researchers view the fundamental nature of science and indeed, reality. These differences led many proto-EDM researchers to work on intelligent tutoring systems and games and at the level of technologies used by individual students, while many proto-LAK researchers worked on collaborative and discussion-based learning environments and more at the university system level. These different choices led to many

researchers not having met, perhaps suggesting why learning analytics researchers founded their own conference rather than joining educational data mining, and why educational data mining researchers (such as the 1st author) largely had not heard of the people who organized the first edition of the LAK conference.

But how do learning analytics researchers and educational data mining researchers (mostly) think differently? To understand this, one possible explanation may date back as far as the 4th century BC, before even the most grizzled veterans of the field were conducting research.

Richard McKeon, a prominent and to many incomprehensible 20th century philosopher, wrote about four philosophical schools of thought dating back to Plato and Aristotle (McKeon 1966): the entitative (Atomist/Democritus), ontological (Platonic), existentialist (Sophist/Protagoras), and essentialist (Aristotelian) schools of thought. Expanded and easier to read explanations of McKeon's writing on these topics can also be found in Watson (1994) and Rich (2018). Each of these schools of thought is associated with preferred "methods" (which, for our purposes, can be thought of as meta-methods or as ways of thinking about methods).

In brief, *reductionism* is the key method of the entitative school. Reductionism can be viewed as an approach towards understanding phenomena that consists of breaking down those phenomena into their constituent components and then analyzing the relationship between those components. Though perhaps reductionism has not captured the hearts and minds of philosophers over the centuries (Democritus is surely less popular than Plato or Aristotle), it has grown to become the core meta-method of many scientific disciplines. Note that one does not need to accept the very strong claim that all phenomena can be reduced to physical processes (i.e. Sachse, 2013), in order to adopt reductionism as a method.

Dialectic is the key method of the ontological school. This school adopts the goal of understanding phenomena as wholes, or understanding systems as systems, where components cannot be properly understood without understanding the whole system. This understanding is often developed through a procedure of developing progressive approximations of a phenomenon or system, through bringing seemingly contradictory perspectives together, and expanding abstract and incomplete understanding towards greater concreteness and completeness.

The terms used by McKeon for the methods of the final two schools (operational and problematic) are a bit less familiar, but the schools of thought themselves are better-known. *Existentialism* views reality as fundamentally individually constructed and therefore asserts that phenomena should be understood as the participants themselves understand them and that these understandings are irreducibly valid (terms such as *philosophical constructionism* and *phenomenological understanding* are sometimes seen). This often leads to methods that are descriptive in nature, as well as a rejection of general-purpose measures (i.e. Strohecker, 1991).

This viewpoint's opposite, *Essentialism*, states that meaning is inherent in the universe. This viewpoint underpins ideas such as a common and universal core of mathematics. It is seen also in perspectives that argue for the "unreasonable effectiveness of data" as justification for rejecting interpretable modeling methods (Halevy et al., 2009), where direct modeling of reality is seen as sufficient and no attempt at theory or explanation is needed (or, indeed, desired).

Our discussion of these issues must admit a debt to a second Richard: Richard "Dick" Buchanan, a key 20th and 21st century thinker on design. For around a decade, Buchanan's design seminar at Carnegie Mellon University discussed these four philosophical schools of thought in terms of approaches to interaction design. Although we as authors are not aware of a formal write-up of Buchanan's ideas around

this, several copies of lecture notes are circulating around the internet¹. Buchanan argued that design in the entitative perspective related a person to an object, design in the ontological perspective related a person to the cosmos, design in the existentialist perspective related people to people, and design in the essentialist perspective related a person to the environment.

One key theme that Buchanan noted in his lectures was that most designers prefer to work in one, or perhaps two, of these perspectives -- and that work from other perspectives often seems confusing or perhaps even intentionally incomprehensible or negative. A clear example of this can be seen in a debate between Greeno (1997) arguing for a ontological “situationalist” perspective on cognitive science and Anderson, Reder, and Simon (1997) arguing for an entitative perspective). Wherein Anderson and colleagues (1997) eventually refer to Greeno’s paradigm as “form”, to their own paradigm as “substance”, and to Greeno’s writing as “situa-babel”. Greeno, for his part, referred to Anderson et al.’s writings as “claims that answer the wrong questions,” a clear sign of different perspectives and priorities.

3.2. The Realization of the Four Paradigms in Learning Analytics

The first attempt to map a perspective like McKeon and Buchanan’s to learning analytics is seen in Siemens and Baker (2012), which suggested that most of the work in educational data mining had a “stronger emphasis on reducing to components and analyzing individual components and relationships between them” (*ibid.*, p. 253) (i.e. entitative, reductionist) whereas most of the work in learning analytics had a “stronger emphasis on understanding systems as wholes, in their full complexity” (*ibid.*, p. 253) (i.e. ontological, dialectical). Indeed, early work in learning analytics argued that the use of reductionism inherently implies a loss of meaning (Atkisson & Wiley, 2011), and the writing of a follow-up to Siemens and Baker’s paper led to a conversation between Siemens and Baker (by email, April 25, 2012) that also illustrated Buchanan’s point around how individuals working in different perspectives see the other perspectives. To quote Siemens, “Question: you mention in the presentation that EDM is reductionist in focus. Is this the message that you want to convey? Reductionism, at least in humanities/social sciences generally has negative connotations. I think labelling EDM as a reductionist approach may paint EDM into a corner where it doesn't want to be! (I'm sure you're aware of the reductionist views, but thought I'd just raise the topic in case you want to revisit the language).” -- the inclusion of this quote is not meant to criticize Siemens (who is very open to other perspectives), but to note how one’s perspective’s terms can become seen as pejorative to members of another community. A similar example is seen in how some statisticians use the term “data mining” to refer to unprincipled, low quality work (i.e. Jensen, 2000).

Based on the paper and discussion between Siemens and Baker, one might reasonably view learning analytics as (starting out as) primarily ontological/dialectical and educational data mining as (starting out as) primarily entitative/reductionist. Figure 1 visualizes the relations between the four paradigms and adjacent research communities.

A few examples of clearly entitative/reductionist work in the early years of EDM include:

- attempts to map out the structure of knowledge domains by reducing the domains to a set of skills which individually items correspond to (Barnes, 2005) and then whether one skill can compensate for another skill within specific items (Pardos et al., 2008).
- attempts to determine which approach is most effective for predicting future student performance within intelligent tutoring systems (Baker et al., 2008; Pavlik et al., 2009)
- attempting to assess the quality of help resources by identifying cases where they are used and looking for immediate positive impact on performance (e.g. Chang et al., 2006).

¹ See for instance, <http://jamin.org/understanding-interaction/> or <https://www.ghostinthepixel.com/?p=319>

A few examples of clearly ontological/dialectical work in just the first year of LAK include:

- The creation of an ontology for multi-level analysis of distributed learning (Suthers & Rosen, 2011)
- Visualizing the shift in interaction patterns between learners over time (Bakharia & Dawson, 2011)
- Sociocultural analysis of the forms of dialogue that emerge between participants in a conference (Ferguson & Shum, 2011)

The Four Paradigms of Learning Analytics

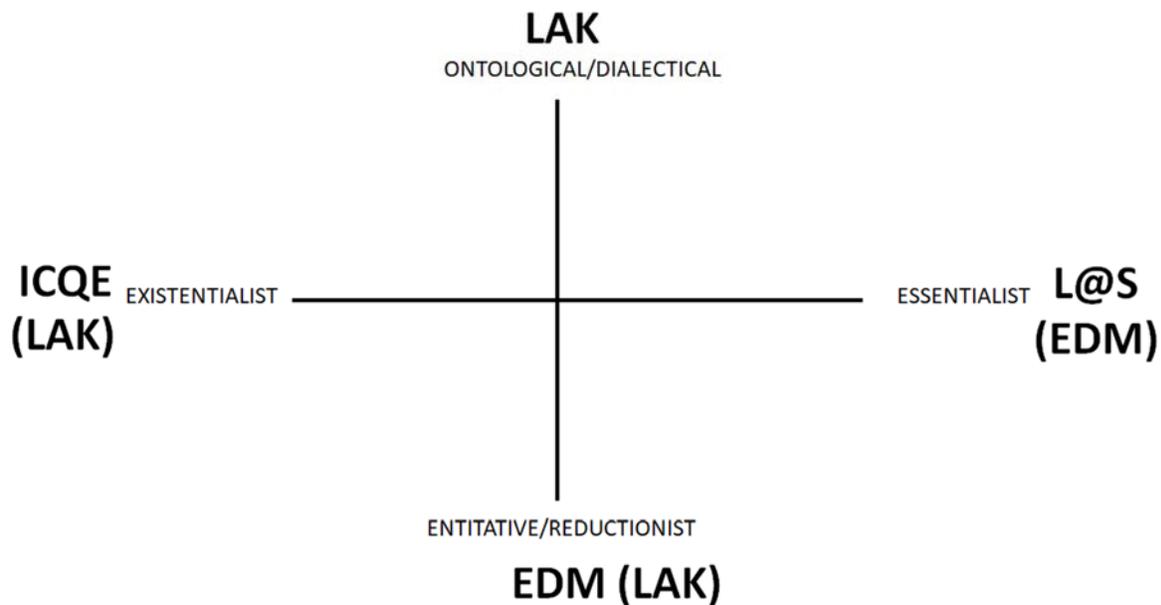


Figure1. A depiction of the four paradigms of learning analytics, and how they are realized in the four learning analytics conferences.

Work from a more essentialist paradigm tended to fare relatively poorly at EDM and LAK in these early years, based on reviewers who expressed negative attitudes towards papers that predicted phenomena without attempting to understand them. A somewhat sarcastic account in Baker (2012), for instance, refers to attempts to predict the final grade in a course using the grades on all of the assignments and examinations. However, the emergence of the ACM Learning @ Scale conference provided a home for more essentialist work.

Early Work in ACM Learning @ Scale, for instance, included:

- An exhaustive examination of when learners in massive online open courses cease watching courses and when interaction is more frequent (Kim et al., 2014)
- An analysis of the characteristics of students who post considerably more frequently than others, including their demographic characteristics, their grades, the number of courses they take, and the content and context of their posts (Huang et al., 2014)

- Clustering the patterns of usage of operators used by students learning an informal programming language, corresponding to the speed with which learners adopt new operators (Yang et al., 2015)

Not all work at ACM Learning @ Scale fit this paradigm – indeed, the second year of this conference (after abandoning a policy explicitly discouraging submission of work involving data from intelligent tutoring systems) had considerable work from authors that generally published at EDM. However, ACM Learning @ Scale presented a venue interested in more essentialist work. With the success of ACM Learning @ Scale, and the popular movement within machine learning towards algorithms that do not attempt to be scrutible (i.e., deep learning; LeCun et al., 2015), there was a movement towards more serious consideration of prediction without comprehension (again, “the unreasonable effectiveness of big data,” Halevy et al., 2009) within EDM, including algorithms that used deep learning to model student knowledge (Khajah et al., 2016), student affect (Botelho et al., 2018), and to grade short-answer responses (Zhang et al., 2016).

This movement was matched by a corresponding movement of entitative work into LAK, with many researchers who had previously published at AIED or EDM beginning to publish their work at LAK. For example:

- Work to model how the perception of ability and the degree of agreement between learners influences the provision of assistance (Choi et al., 2019)
- Work to understand how changes in the design of learning games concretely impact learner choices and performance (Harpstead et al., 2019)
- An analysis of stopout in homework assignments and how it relates to student self-confidence (Botelho et al., 2019)

What remained generally missing was the fourth, existentialist perspective. Work in this tradition was seen occasionally in the broader learning analytics space, with papers attempting to use analytics and data to capture learning in open-ended and constructionist (self-directed) tasks (Blikstein, 2011; Worsley & Blikstein, 2015), papers attempting to capture the range of different ways that different students approach a problem (Martin et al., 2015), papers using methods like clustering to gain understanding of deeper differences between individuals (Amershi & Conati, 2009), and epistemic network analysis of the difference in patterns of discourse between college and high school students working through the same curriculum (Shaffer & Ruis, 2017). Existential researchers have until this point generally not found a strong home in any of the existing learning analytics communities – for example, Berland, Baker, & Blikstein’s manifesto on the applications of learning analytics to constructionism appeared in *Technology, Knowledge, and Learning* (2014) rather than a mainstream learning analytics venue, and after several early papers in learning analytics venues, Berland and Blikstein have mostly moved to other publication venues in recent years.

Quantitative ethnography researchers, by contrast, responded to the lack of a learning analytics venue that had a prominent place for their work (although epistemic network analysis has been successful at CSCL) by founding their own conference, the International Conference on Quantitative Ethnography. The remit of this conference is more broad than learning or education, and part of the goal of establishing the conference was to broaden the community both in terms of domain areas and methods, but 78% of full papers in the first year of the conference nonetheless were in these domains.

Early work in the first year of ICQE included:

- Analyses of differences in pre-service teachers' diagnostic argumentations (Bauer et al., 2019)
- The use of Epistemic Network Analysis to represent shifts in identity over time among long-term members of an online game community (Barany & Foster, 2019)
- Understanding the differences in processes of how students regulate their collaborative learning when faced with either motivational or comprehension-related problems (Melzner et al., 2019)

Interestingly, research on constructionist learning environments was absent from the first year of the International Conference on Quantitative Ethnography, despite Madison (the city where the conference was held) being a center for constructionist learning analytics research. This may be due in part to the methodological focus of quantitative ethnography – the use of epistemic network analysis and related methods – which represents one take on how to do existentialist work with large-scale educational data but is clearly not the only such method. While some constructionist learning analytics research has involved network analysis methods (i.e. Martin et al., 2015), much of the work in these contexts has involved other approaches. It remains to be seen where constructionist learning analytics researchers will find their intellectual home in the years to come.

Very recently, there has been a shift where more existential work has a greater place in the learning analytics conference and journal. One development along these lines is the recent publication in the Journal of Learning Analytics of a special issue on human-centered learning analytics, which called in its introduction for work to “acknowledge that understandings of reality vary in different contexts... a shift here away from the primarily quantitative focus of LA toward rich accounts that help to uncover why and how analytics are used, and why they may be misused or ignored” (Buckingham Shum et al., 2019, p. 6). Correspondingly, a massive expansion of papers using epistemic network analysis methods appeared in the LAK conference in 2020 (i.e. Chen et al., 2020a; Ferreira et al., 2020; Saint et al., 2020; Swiecki & Shaffer, 2020; Uzir et al., 2020; Whitelock-Wainwright et al., 2020).

Overall, then, each of the four paradigms has a primary home in one of the four learning analytics communities: reductionism at EDM, dialectic at LAK, essentialism at L@S, and existentialism at ICQE. Recently, EDM has broadened towards including essentialism, and LAK has broadened to include both reductionist and existential work. These alignments continue to shift, and it remains to be seen where each of these communities ends up.

4. The Key Place of Each of the Four Paradigms in Learning Analytics Practice

When a field in the sciences contains more than one competing intellectual paradigm, there is often a push for one of the fields to “win” and take over the field entirely, as depicted in Kuhn’s (1962) classic book *The Structure of Scientific Revolutions*. Although Kuhn explicitly disavowed the application of his theory to the social sciences, psychology for instance has seen considerable discussion of his ideas over the span of decades (Coleman & Salamon, 1988). Certainly, the type of rhetoric seen in Anderson, Reder, and Simon (1997) and Greeno (1997) demonstrate the sort of active debate between members of different paradigms about which paradigm is *right*. The speciation of conferences in the broader field of learning analytics also demonstrates the desire for researchers to find an intellectual home that matches their own way of looking at the world – as does, perhaps, the migration of work from EDM (entitative) to LAK (dialectical), as the intellectual ideas of essentialism migrated into EDM.

What we hope with this article, as researchers/practitioners mostly working in different paradigms, is to argue that different modes of thought are natural, are positive, and are good for the field. The complex challenges that learning analytics poses to us as researchers and practitioners (cf. Baker, 2019; Pelanek,

2020) are too large to be entirely resolved by any of these four paradigms. The opportunities that learning analytics holds for the world of education are too important to be squandered in internecine squabbles.

However, we would argue that an attitude of “live and let live,” where the various conferences go their separate ways is not desirable (and fortunately, is not happening, as our discussion above indicates). The natural tendency is for a conference and community to select research problems that appeal to its intellectual tendencies and focus on these problems. Perhaps, the best example of this in the broader field of learning analytics is the intense focus at the EDM conference on predicting next problem correctness and/or inferring knowledge. In 2019, the EDM conference had 28 articles on this topic!

Instead, we suggest that there needs to be greater collaboration across researchers from different intellectual paradigms – a move towards *inter-paradigmatic* work in addition to the inter-disciplinarity that already characterizes our field. As Pavlik & Toth (2010, p. 105) note, in talking about a different set of perspectives within the intelligent tutoring system community of practice, “Essentially, while there may be some differences between the perspectives... in this situation the perspectives complement and strengthen each other.”

In this spirit, we would like to offer a few possible cases where we believe that one paradigm’s methods may be highly useful for addressing another paradigm’s challenges.

The first of our possibilities is already happening. Inscrutable models that fit data better, from an essentialist paradigm, are migrating into EDM. The reason for this is that better prediction is, in many cases, a good thing in itself. While predictive accuracy is not the only goal for a model, there are good reasons to prefer a model that achieves an AUC ROC of 0.75 to a model that achieves an AUC ROC of 0.65. Inscrutable, highly-predictive models have had transformative impacts on our daily life, in areas such as voice recognition. The key is, perhaps, to identify the cases where higher predictive accuracy is the top goal, and have good ways to check for and mitigate the other validity issues that may come up when accuracy is optimized. Yeung and Yeung’s (2018) shows an excellent example of this, taking a deep learning model for student knowledge estimation, identifying validity issues, and using regularization to fix these issues without reducing model predictive accuracy. Integrating the predictive accuracy benefits of essentialism with a careful entitative consideration of validity may be beneficial for solving the problems that entitative and dialectical learning analytics practitioners are trying to solve.

The integration of essentialist high predictive power with entitative component-by-component consideration of validity may produce measures that can be highly useful to other paradigms. Entitative work within EDM has proven useful in a range of “discovery with models” analyses, where a model of a phenomenon is used to more deeply understand that phenomenon or, indeed, other phenomena (Baker & Yacef, 2009). Discovery with models has been a prominent method within EDM over the years – 19% of papers in Baker and Yacef’s (2009) original survey and 18% of full papers in EDM in 2019. LAK, in its study of systems, has often relied on relatively simple measures of constructs such as engagement, with resultant challenges in deciding how to define those metrics (i.e. Kovanovic et al., 2016). Substituting more predictive and more thoroughly-validated models into these analyses may result in richer and more trustworthy findings. Similarly, much of the work in quantitative ethnography has inputted relatively simple metrics into complex epistemic network analyses, or has relied upon either exhaustive human coding or the use of regular expression matching to generate data to input into epistemic network analyses (e.g. Cai et al., 2019). Again, substituting highly predictive and thoroughly-validated models into epistemic network analyses may enable deep consideration of the inter-relationships between the complex types of constructs that EDM and ACM L@S work can now measure (e.g., automatic coding of discourse, Ferreira et al., 2020).

The ability of entitative work to detect complex constructs has another potential benefit for work in other paradigms. Much existentialist work involves qualitative, deep analysis of specific cases, which is an arduous process (Sherin et al., 2018). One possible approach is to use visualization to speed this process (Shaffer & Ruis, 2017; Sherin et al., 2018). However, the very collection of qualitative data can often be a very in-depth and time-consuming process. Using automated models to drive both the collection and analysis of qualitative data may have major benefits – if we can identify the specific cases we are particularly interested in, we can focus analysis on those cases. More powerfully, if we can identify specific cases we are interested in *as they are occurring*, we can focus qualitative data collection -- such as interviews and observations -- on those cases. An example of this is seen in ongoing work where detectors of shifts in student affect are used to trigger alerts to a qualitative field researcher, who then goes and interviews the student within one minute of the shift occurring (Baker et al., in preparation).

There are several ways that work from an existentialist perspective could support work in other paradigms. Quantitative ethnography and other existentialist methods have considerable potential to support understanding of the phenomena that are studied by machine learned models in other paradigms. Comparing the relationships between behaviors (i.e. comparing epistemic networks) between cases where a phenomenon is seen and cases where a phenomenon is not seen has the potential to increase understanding of why that phenomenon occurs (e.g. Karumbaiah et al., 2019), leading to better feature engineering for detectors of that phenomenon. Phenomenological attempts to understand the reasoning human coders use when they identify a construct may also lead to better detection of that construct. An example of this is seen in work by Paquette and colleagues (2014), who used knowledge engineering methods to understand how human coders identify gaming the system, and then used it as the basis of a detector of that construct that transferred between learning systems with no re-training (Paquette & Baker, 2019).

Ethnographic and other existentialist work may have another important application in learning analytics – not in the step of modeling but in the step of using those models. As noted above, ethnographic work likely has an important role in designing interventions – whether reports or automated interventions – that teachers, students, and other stakeholders find useful, useable, and desirable (cf. Buckingham Shum et al., 2019; Holstein et al., 2019). While this may generally take the form of collaboration between dialectical or entitative learning analytics and human-computer interaction practitioners or traditional ethnographers, quantitative ethnography has the potential to develop models that better align with stakeholders’ thinking and which are therefore more desirable to them. There is often skepticism voiced by users about how and whether models work (Rientes et al., 2018) – better alignment to how humans think may be a step towards addressing the barriers to adoption that challenge learning analytics applications.

Going in the other direction, learning analytics has the potential to offer system-level understanding in quantitative ethnography. As an existentialist methodological approach, ethnography offers rich insights into individual experiences which are deeply contextualized. While it can offer invaluable insights for building models and tools that are well-aligned with stakeholders’ thinking, it is difficult within this type of method to consider broader system level perspectives. For learning analytics with a strong ontological perspective, a system level understanding is expected. An example of a productive direction of future research is to examine how rich ethnographic accounts of stakeholder expectations from and experiences with learning analytics (Pontual Falcão et al., 2020) can be incorporated into models that can simulate systemic impact of relevant policies and strategies (Dawson et al., 2019). Another related example of adding a system perspective to ethnographic accounts in learning analytics is Social and Epistemic Network Signature (SENS) (Gašević et al., 2019; Swiecki & Shaffer, 2020). SENS is an approach that

integrates social network analysis (ontological) and epistemic network analysis (existentialist) to support analysis of the social and cognitive dimensions of networked and collaborative learning.

5. Conclusion

In this article, we have discussed the interrelationships and core differences between four communities interested in the role of big data and data science for education: learning analytics, educational data mining, learning at scale, and quantitative ethnography.

We propose that the different work seen in these four communities can be explained in significant part by four different philosophical paradigms dating back to the beginnings of philosophy in Ancient Greece, adopting and adapting a framework proposed by Richard McKeon and Richard Buchanan. While the four conferences do not map exactly to these four paradigms, the framework nonetheless has a surprising degree of fit to the different work seen in these four communities of learning analytics.

We call for a better understanding, within each of us, of how our philosophical stances impact our research and practice. By understanding ourselves, we can understand the deep perspectives that lead to the specific choices we make, and seek analogies in the history of other work coming from that philosophical tradition. By understanding the philosophical stances that our colleagues in our broader community bring to bear on their own research and practice, we can better understand why they make the choices they make, and how these choices emerge from deep (if sometimes unspoken) philosophical commitments rather than from ignorance, laziness, or foolishness.

Perhaps even more important, by understanding our colleagues' philosophical stances, we may be able to better see what they may be able to see that we cannot see; we may be able to better learn from them; and we may be more able to craft research projects that take full advantage of what we each have to contribute to our field and to the learners we as communities strive to serve. This article attempts to propose new connections and projects that may become possible by working together. There are many more possibilities for connection and collaboration, which we have not yet envisioned. Ultimately, the challenges our field hopes to solve cannot be solved by a single perspective or toolkit. We must work together to solve them.

References

- Amershi, S., & Conati, C. (2009). Combining unsupervised and supervised classification to build user models for exploratory learning environments. *Journal of Educational Data Mining*, 1(1), 18-71. <https://doi.org/10.5281/zenodo.3554659>
- An, P., Bakker, S., Ordanovski, S., Taconis, R., Paffen, C. L., & Eggen, B. (2019, May). Unobtrusively enhancing reflection-in-action of teachers through spatially distributed ambient information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (paper no. 91). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300321>
- Anderson, J. R., Reder, L. M., & Simon, H. A. (1997). Situative versus cognitive perspectives: Form versus substance. *Educational Researcher*, 26(1), 18-21. <https://doi.org/10.3102%2F0013189X026001018>
- Atkisson, M., & Wiley, D. (2011). Learning analytics as interpretive practice: applying Westerman to educational intervention. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 117-121). Association for Computing Machinery. <https://doi.org/10.1145/2090116.2090133>
- Baker, R.S. (2012) *Big data and education* (1st ed.). Teachers College, Columbia University.
- Baker, R.S. (2019) Challenges for the future of educational data mining: The Baker learning analytics prizes. *Journal of Educational Data Mining*, 11(1), 1-17. <https://doi.org/10.5281/zenodo.3554745>
- Baker, R.S.J.d., Corbett, A.T., Aleven, V. (2008) Improving contextual models of guessing and slipping with a truncated training set. In R. de Baker, T. Barnes, & J. Beck (Eds.), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 67-76). International Educational Data Mining Society.
- Baker, R.S.J.d., 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. <https://doi.org/10.5281/zenodo.3554657>
- Bakharia, A., & Dawson, S. (2011). SNAPP: a bird's-eye view of temporal participant interaction. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 168-173). Association for Computing Machinery. <https://doi.org/10.1145/2090116.2090144>
- Barany, A., & Foster, A. (2019). Examining identity exploration in a video game participatory culture. In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in Quantitative Ethnography* (pp. 3-13). Springer. https://doi.org/10.1007/978-3-030-33232-7_1
- Barnes, T. (2005). The Q-matrix method: Mining student response data for knowledge. In J. E. Beck (Ed.), *Proceedings of AAAI 2005 Educational Data Mining Workshop* (pp. 39-46). Association for the Advancement of Artificial Intelligence.
- Bauer, E., Sailer, M., Kiesewetter, J., Schulz, C., Pfeiffer, J., Gurevych, I., Fischer, M., & Fischer, F. (2019). Using ENA to analyze pre-service teachers' diagnostic argumentations: A conceptual framework and initial applications. In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in Quantitative Ethnography* (pp. 14-25). Springer. https://doi.org/10.1007/978-3-030-33232-7_2
- Berland, M., Baker, R.S., Blikstein, P. (2014) Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge, and Learning*, 19, 205-220. <https://doi.org/10.1007/s10758-014-9223-7>
- Blikstein, P. (2011). Using learning analytics to assess students' behavior in open-ended programming tasks. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 110-116). Association for Computing Machinery. <https://doi.org/10.1145/2090116.2090132>
- Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, 10(4), 405-418. <https://doi.org/10.1109/TLT.2017.2740172>
- Botelho, A.F., Baker, R., Ocumpaugh, J., Heffernan, N. (2018) Studying affect dynamics and chronometry using sensor-free detectors. In K. E. Boyer, M. Yudelson (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining* (pp. 157-166). International Educational Data Mining Society.
- Botelho, A. F., Varatharaj, A., Inwegen, E. G. V., & Heffernan, N. T. (2019). Refusing to try: Characterizing early stopout on student assignments. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 391-400). Association for Computing Machinery. <https://doi.org/10.1145/3303772.3303806>
- Bowers, A. J., Sprott, R., & Taff, S. A. (2012). Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity, and specificity. *The High School Journal*, 96(2), 77-100. <https://doi.org/10.1353/hsj.2013.0000>
- Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, 6(2), 1-9. <https://doi.org/10.18608/jla.2019.62.1>
- Buckingham Shum, S., & Luckin, R. (2019). Learning analytics and AI: Politics, pedagogy and practices. *British Journal of Educational Technology*, 50(6), 2785-2793. <https://doi.org/10.1111/bjet.12880>
- Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D. W., Hu, X., & Graesser, A. C. (2019). nCoder+: A semantic tool for improving recall of nCoder coding. In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in Quantitative Ethnography* (pp. 41-54). Springer. https://doi.org/10.1007/978-3-030-33232-7_4
- Campbell, J. P., DeBlois, P. B., & Oblinger, D. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*, 42(4), 40-57.
- Chang, K. M., Beck, J. E., Mostow, J., & Corbett, A. (2006). Does help help? A Bayes net approach to modeling tutor interventions. In *Proceedings of the 21st Annual Meeting of the American Association for Artificial Intelligence* (pp. 41-46). American Association for Artificial Intelligence.
- 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>
- Chaturapruek, S., Dee, T. S., Johari, R., Kizilcec, R. F., & Stevens, M. L. (2018). How a data-driven course planning tool affects college students' GPA: Evidence from two field experiments. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (pp. 63:1-63:10). Association for Computing Machinery. <https://doi.org/10.1145/3231644.3231668>
- Chen, G., Rolim, V., Mello, R. F., & Gašević, D. (2020a). Let's shine together! a comparative study between learning analytics and educational data mining. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 544-553). Association for Computing Machinery. <https://doi.org/10.1145/3375462.3375500>
- Chen, X., Zou, D., Cheng, G., & Xie, H. (2020b). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of *Computers & Education*. *Computers & Education*, 151, 103855
- Choi, H., Dowell, N., Brooks, C., & Teasley, S. (2019). Social comparison in MOOCs: Perceived SES, opinion, and message

- formality. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 160-169). Association for Computing Machinery. <https://doi.org/10.1145/3303772.3303773>
- Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). Association for Computing Machinery. <https://doi.org/10.1145/2330601.2330636>
- Coleman, S. R., & Salamon, R. (1988). Kuhn's *Structure of Scientific Revolutions* in the psychological journal literature, 1969-1983: A descriptive study. *The Journal of Mind and Behavior*, 9(4), 415-446.
- Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S., Nelson, K., Alexander, S., Lockyer, L., Kennedy, G., Corrin, L., & Fisher, J. (2015). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*. Office of Learning and Teaching, Australian Government. http://he-analytics.com/wp-content/uploads/SP13-3249_Master17Aug2015-web.pdf
- Corrin, L., & de Barba, P. (2014). Exploring students' interpretation of feedback delivered through learning analytics dashboards. In *Proceedings of the ASCILITE 2014 Conference* (pp. 629–633). Australasian Society for Computers in Learning in Tertiary Education.
- Dawson, S., Joksimovic, S., Poquet, O., & Siemens, G. (2019). Increasing the impact of learning analytics. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 446-455). Association for Computing Machinery. <https://doi.org/10.1145/3303772.3303784>
- 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). Association for Computing Machinery. <https://doi.org/10.1145/3170358.3170375>
- Essa, A. (2016). A possible future for next generation adaptive learning systems. *Smart Learning Environments*, 3(16), 1-24. <https://doi.org/10.1186/s40561-016-0038-y>
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5), 304–317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Ferguson, R., & Shum, S. B. (2011). Learning analytics to identify exploratory dialogue within synchronous text chat. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 99-103). Association for Computing Machinery. <https://doi.org/10.1504/IJTEL.2012.051816>
- Ferreira, M., Rolim, V., Mello, R. F., Lins, R. D., Chen, G., & Gašević, D. (2020). Towards automatic content analysis of social presence in transcripts of online discussions. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 141-150). Association for Computing Machinery. <https://doi.org/10.1145/3375462.3375495>
- Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, 92, 562-577. <https://doi.org/10.1016/j.chb.2018.07.003>
- 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>
- Google Scholar. (2020). Educational technology [Citation count]. Retrieved from https://scholar.google.ca/citations?view_op=top_venues&hl=en&vq=soc_educationaltechnology
- Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17. <https://journals.sagepub.com/doi/10.3102/0013189X026001005>
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, 15(3), 42–57.
- Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2), 8-12. <https://doi.org/10.1109/MIS.2009.36>
- Harpstead, E., Richey, J. E., Nguyen, H., & McLaren, B. M. (2019). Exploring the subtleties of agency and indirect control in digital learning games. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 121-129). Association for Computing Machinery. <https://doi.org/10.1145/3303772.3303797>
- Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, 6(2), 27–52. <https://doi.org/10.18608/jla.2019.62.3>
- Huang, J., Dasgupta, A., Ghosh, A., Manning, J., & Sanders, M. (2014). Superposter behavior in MOOC forums. In *Proceedings of the First ACM conference on Learning @ Scale Conference* (pp. 117-126). Association for Computing Machinery. <https://doi.org/10.1145/2556325.2566249>
- Jacobson, M. J., Kapur, M., & Reimann, P. (2016). Conceptualizing debates in learning and educational research: Toward a complex systems conceptual framework of learning. *Educational Psychologist*, 51(2), 210–218. <https://doi.org/10.1080/00461520.2016.1166963>

- Jensen, D. (2000). Data snooping, dredging and fishing: The dark side of data mining a SIGKDD99 panel report. *SIGKDD Explorations*, 1(2), 52-54. <https://doi.org/10.1145/846183.846195>
- Jovanovic, J., Gasevic, D., Brooks, C., Devedzic, V., Hatala, M., Eap, T., & Richards, G. (2008). LOCO-Analyst: Semantic web technologies in learning content usage analysis. *International Journal of Continuing Engineering Education and Life Long Learning*, 18(1), 54–76. <https://doi.org/10.1504/IJCEELL.2008.016076>
- Karumbaiah, S., Baker, R.S., Barany, A., Shute, V. (2019) Using epistemic networks with automated codes to understand why players quit levels in a learning game. In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in Quantitative Ethnography* (pp. 106-116). Springer. https://doi.org/10.1007/978-3-030-33232-7_9
- Khajah, M., Lindsey, R. V., & Mozer, M. C. (2016) How deep is knowledge tracing?. In T. Barnes, M. Chi, M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 94-101). International Educational Data Mining Society.
- Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference* (pp. 31-40). Association for Computing Machinery. <https://doi.org/10.1145/2556325.2566237>
- Kovanovic, V., Gasevic, D., Dawson, S., Joksimovic, S., Baker, R.S., & Hatala, M. (2016) Does time-on-task estimation matter? Implications on validity of learning analytics findings. *Journal of Learning Analytics*, 2(3), 81-110. <https://doi.org/10.18608/jla.2015.23.6>
- Kuhn, T.S. (1962). *The structure of scientific revolutions* (1st ed.). University of Chicago Press.
- Lang, C., Siemens, G., Wise, A., & Gašević, D. (2017). *Handbook of learning analytics* (1st ed.). Society for Learning Analytics and Research. <https://doi.org/10.18608/hla17>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Li, H., Gobert, J., Dickler, R., & Moussavi, R. (2018). The impact of multiple real-time scaffolding experiences on science inquiry practices. In Nkambou R., Azevedo R., & Vassileva J. (Eds.), *Intelligent Tutoring Systems: 14th International Conference, Proceedings* (pp. 99-109). Springer. https://doi.org/10.1007/978-3-319-91464-0_10
- Lingard, B., & Lewis, S. (2016). Globalisation of the Anglo-American approach to top-down, test-based educational accountability. In G. T. L. Brown, L R. Harris (Eds.), *Handbook of human and social conditions in assessment* (1st ed., pp. 387-403).
- Long, P. D., Siemens, G., Conole, G., & Gašević, D. (Eds.). (2011). *LAK'11: Proceedings of the 1st international conference on learning analytics and knowledge*. Association for Computing Machinery.
- Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, 47, 90–97. <https://doi.org/10.1016/j.chb.2014.07.013>
- Lynch, C. F. (2017). Who prophets from big data in education? New insights and new challenges. *Theory and Research in Education*, 15(3), 249-271. <https://doi.org/10.1177/1477878517738448>
- Macfadyen, L. P., Dawson, S., Pardo, A., & Gasevic, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. *Research & Practice in Assessment*, 9(Winter), 17–28.
- Martin, T., Petrick Smith, C., Forsgren, N., Aghababayan, A., Janisiewicz, P., & Baker, S. (2015). Learning fractions by splitting: Using learning analytics to illuminate the development of mathematical understanding. *Journal of the Learning Sciences*, 24(4), 593-637. <https://doi.org/10.1080/10508406.2015.1078244>
- McKeon, R. (1966). Philosophic semantics and philosophic inquiry. In Z. K. McKeon (Ed.), *Freedom and history and other essays: An introduction to the thought of Richard McKeon* (pp. 242-256). The University of Chicago Press.
- Melzner, N., Greisel, M., Dresel, M., & Kollar, I. (2019). Using Process Mining (PM) and Epistemic Network Analysis (ENA) for comparing processes of collaborative problem regulation. In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in Quantitative Ethnography* (pp. 154-164). Springer. https://doi.org/10.1007/978-3-030-33232-7_13
- Paquette, L., & Baker, R. S. (2019). Comparing machine learning to knowledge engineering for student behavior modeling: a case study in gaming the system. *Interactive Learning Environments*, 27(5-6), 585-597. <https://doi.org/10.1080/10494820.2019.1610450>
- Paquette, L., de Carvalho, A. M. J. A., Baker, R.S. (2014) Towards understanding expert coding of student disengagement in online learning. In *Proceedings of the 36th Annual Cognitive Science Conference* (pp. 1126-1131). Cognitive Science Society.
- Pardos, Z., Heffernan, N., Ruiz, C., & Beck, J. (2008). The composition effect: Conjunctive or compensatory? an analysis of multi-skill math questions in ITS. In R.S.J.d. Baker, T. Barnes, J. E. Beck (Eds.), *Proceedings of the 1st International Conference on Educational Data Mining* (pp. 147-156). International Educational Data Mining Society.
- Pavlik Jr, P. I., Cen, H., & Koedinger, K. R. (2009). Learning factors transfer analysis: Using learning curve analysis to

- automatically generate domain models. In T. Barnes, M. Desmararis, C. Romero, & S. Ventura (Eds.), *Proceedings of the 2nd International Conference on Educational Data Mining* (pp. 121-130). International Educational Data Mining Society.
- Pavlik, P., & Toth, J. (2010). How to build bridges between intelligent tutoring system subfields of research. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Intelligent Tutoring Systems: 10th International Conference, Proceedings Part II* (pp. 103-112). Springer. https://doi.org/10.1007/978-3-642-13437-1_11
- Pelánek, R. (2020). Learning analytics challenges: trade-offs, methodology, scalability. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 554-558). Association for Computing Machinery. <https://doi.org/10.1145/3375462.3375463>
- Pontual Falcão, T., Ferreira Mello, R., Rodrigues, R. L., Diniz, J. R. B., Tsai, Y. S., & Gašević, D. (2020). Perceptions and expectations about learning analytics from a brazilian higher education institution. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 240-249). <https://doi.org/10.1145/3375462.3375478>
- Rich, M. H. (2018). McKeon's semantics of communication: A pragmatic exploration of the communicative arts. *Empedocles: European Journal for the Philosophy of Communication*, 9(1), 5-24. https://doi.org/10.1386/ejpc.9.1.5_1
- Rienties, B., Herodotou, C., Olney, T., Schencks, M., & Borooowa, A. (2018). Making sense of learning analytics dashboards: A technology acceptance perspective of 95 teachers. *International Review of Research in Open and Distributed Learning*, 19(5), 186-202. <https://doi.org/10.19173/irrodl.v19i5.3493>
- Ritter, S., Yudelson, M., Fancsali, S. E., & Berman, S. R. (2016). How mastery learning works at scale. In *Proceedings of the Third ACM Conference on Learning @ Scale* (pp. 71-79). Association for Computing Machinery. <https://doi.org/10.1145/2876034.2876039>
- Sachse, C. (2013). *Reductionism in the philosophy of science* (Vol. 11). Walter de Gruyter.
- Saint, J., Gašević, D., Matcha, W., Uzir, N. A. A., & Pardo, A. (2020). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 402-411). <https://doi.org/10.1145/3375462.3375487>
- Sales, A., Botelho, A. F., Patikorn, T., & Heffernan, N. T. (2018). Using big data to sharpen design-based inference in A/B tests. In K. E. Boyer, M. Yudelson (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining* (pp. 479-485). International Educational Data Mining Society.
- Sclater, N., & Bailey, P. (2015). *Code of practice for learning analytics*. Jisc. <https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics>
- Settles, B., & Meeder, B. (2016). A trainable spaced repetition model for language learning. In K. Erk, & N. A. Smith (Eds.), *Proceedings of the 54th annual meeting of the association for computational linguistics (Volume 1: Long papers)* (pp. 1848-1858). <https://doi.org/10.18653/v1/P16-1174>
- Shaffer, D., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 175-187). <https://doi.org/10.18608/hla17.015>
- Sherin, B., Kersting, N., Berland, M. (2018) Learning analytics in support of qualitative analysis. In J. Kay, & R. Luckin (Eds.), *Proceedings of the 13th International Conference of the Learning Sciences* (pp. 464-471). International Society of the Learning Sciences.
- 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. (2014). The journal of learning analytics: Supporting and promoting learning analytics research. *Journal of Learning Analytics*, 1(1), 3-5. <https://doi.org/10.18608/jla.2014.11.2>
- Singh, R. P. (2018). Learning analytics: Potential, protection, and privacy in the educational system. In M. K. Singh, Z. Zerihun, N. Singh (Eds.), *Impact of learning analytics on curriculum design and student performance* (pp. 1-18). IGI Global. <https://doi.org/10.4018/978-1-5225-5369-4.ch001>
- Steiner, C., Kickmeier-Rust, M., & Türker, M. A. (2014). *Deliverable D3.1 review article about LA and EDM approaches*. European Commission LEA's Box Project. <http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D3.1.pdf>
- Strohecker, C. (1991). Elucidating styles of thinking about topology through thinking about knots. In I Harel, & S. Papert (Eds.), *Constructionism*, (pp. 215-234). Ablex.
- Suthers, D., & Rosen, D. (2011). A unified framework for multi-level analysis of distributed learning. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 64-74). Association for Computing Machinery. <https://doi.org/10.1145/2090116.2090124>
- Swiecki, Z., & Shaffer, D. W. (2020). iSENS: an integrated approach to combining epistemic and social network analyses. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 305-313). Association for Computing Machinery. <https://doi.org/10.1145/3375462.3375505>

- Tsai, Y.-S., Gašević, D., Whitelock-Wainwright, A., Muñoz-Merino, P. J., Moreno-Marcos, P. M., Rubio Fernández, A., Delgado Kloos, C., Scheffel, M., Jivet, I., Drachsler, H., Tammets, K., Ruiz Calleja, A., & Kollom, K. (2018). *SHEILA: Supporting higher education to integrate learning analytics*. The University of Edinburgh. <http://sheilaproject.eu/wp-content/uploads/2018/11/SHEILA-research-report.pdf>
- Tsai, Y.-S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2019). Complexity leadership in learning analytics: Drivers, challenges and opportunities. *British Journal of Educational Technology*, 50(6), 2839–2854. <https://doi.org/10.1111/bjet.12846>
- Uzir, N. A. A., Gašević, D., Jovanović, J., Matcha, W., Lim, L. A., & Fudge, A. (2020). Analytics of time management and learning strategies for effective online learning in blended environments. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 392-401). Association for Computing Machinery. <https://doi.org/10.1145/3375462.3375493>
- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89, 98-110. <https://doi.org/10.1016/j.chb.2018.07.027>
- Watson, W. (1994). McKeon's semantic schema. *Philosophy & Rhetoric*, 27(2), 85-103.
- Weller, M. (2020). *25 years of ed tech*. Athabasca University Press. <https://doi.org/10.15215/aupress/9781771993050.01>
- Whitelock-Wainwright, A., Tsai, Y. S., Lyons, K., Kaliff, S., Bryant, M., Ryan, K., & Gašević, D. (2020). Disciplinary differences in blended learning design: A network analytic study. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 579-588). <https://doi.org/10.1145/3375462.3375488>
- Wise, A., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, 6(2), 53–69. <https://doi.org/10.18608/jla.2019.62.4>
- Worsley, M., & Blikstein, P. (2015). Leveraging multimodal learning analytics to differentiate student learning strategies. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 360-367). Association for Computing Machinery. <https://doi.org/10.1145/2723576.2723624>
- Yang, S., Domeniconi, C., Revelle, M., Sweeney, M., Gelman, B. U., Beckley, C., & Johri, A. (2015). Uncovering trajectories of informal learning in large online communities of creators. In *Proceedings of the Second ACM Conference on Learning @ Scale* (pp. 131-140). Association for Computing Machinery. <https://doi.org/10.1145/2724660.2724674>
- Yeung, C. K., & Yeung, D. Y. (2018). Addressing two problems in deep knowledge tracing via prediction-consistent regularization. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (pp. 1-10). Association for Computing Machinery. <https://doi.org/10.1145/3231644.3231647>
- Zhang, Y., Shah, R., & Chi, M. (2016). Deep learning+ student modeling+ clustering: A recipe for effective automatic short answer grading. In T. Barnes, M. Chi, M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 562-567). International Educational Data Mining Society.