INTERNATIONAL JOURNAL OF MULTIPLE RESEARCH APPROACHES, 2018
VOL. 10, NO. 1, 102–111
https://doi.org/10.29034/ijmra.v10n1a6

Dialectical Publishing

Learning Analytics as an Educational Research Approach

Donggil Song

Department of Computer Science, Sam Houston State University, TX, USA

ABSTRACT

This article reviews Learning Analytics (LA) to understand the nature and characteristics of LA better, specifically, how this research approach is utilized in educational studies. Educational studies utilizing the LA approach have made it possible to look at the effect of changes related to diverse learning variables over time. These variables span across learners, instructors, and the relationship of those changes to the performance of learners. Currently, there is a growing body of literature focusing on the use of LA, such as educational data mining, data visualization, and numerical modeling. Rapid advancements in the capacity of computer processing have showcased the potential of LA, which encompasses big data processing and the immense quantities of the data analysis. This article contemplates challenges and opportunities for the use of LA as a research approach. Implications of previous studies and traditions for the conceptualization and conduct of LA research are also discussed. Particularly, this study focuses on what brings LA to the research field, what research questions LA have answered, and how the approach is implemented. Additional methodological issues are also discussed.

KEYWORDS

Educational data mining; learning analytics; log data analysis; research approach

The motivation for this article derived from the fact that in-depth contemplation is required for methodological frameworks of Learning Analytics (LA) to gain acceptance in the academic community. A search of relevant literature did not reveal robust consideration of the added value of LA in the research approach domain, more specifically, in the educational research methods field. Consequently, there is a need to supply educational researchers with an accredited overview of LA. This article aims to fill that gap as well as to carry out a review of LA studies to contribute towards a documentation of the LA research approach so far. This review includes what brings LA to the educational research field, what research questions LA have answered, and how the approach is implemented. Along with the additional methodological issues, this review also includes an investigation of LA that captures the strengths and weaknesses in data analysis and the identification of purposes of these previous studies, and thus, hopefully, motivate the research community to reconceptualize LA as a research approach for further research.

Introduction of Learning Analytics

Before addressing LA, the last term, *analytics*, needs to be delineated first. In general, analytics refers to "a generic set of techniques and algorithms that have been used for quite some time" in some domains (Pardo, 2014, p. 16). The first term, *learning*, specifies the field or topic of research in analytics. Thus, LA can be defined as a set of techniques and algorithms that are used in the learning-related domain. Per the 1st International Conference on Learning Analytics and Knowledge (2010), LA is defined 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" (para. 6). Although there is no agreement on a standard definition of LA, this definition was used in most studies reviewed in the present article. However, the afore-

CONTACT Donggil Song ✉ song@shsu.edu

mentioned definition might have a wide scope of research concepts to a degree. It is difficult to differentiate the concept of LA from the existing research approaches in education because learning science and instructional technology researchers have already measured, collected, analyzed, and reported data about learners and their contexts without using LA. Thus, we need a further narrowed-down definition for the term, LA.

Educational researchers have characterized LA as a research field, discipline, or approach, with certain types of techniques. Knight, Shum, and Littleton (2014) considered LA as an emerging "research field and design discipline" (p. 2), specifically, a field of the educational research that uses computational techniques to capture and to analyze learning-related data. Scheffel, Drachsler, Stoyanov, and Specht (2014) also stipulated LA as a multi-disciplinary research field that is built upon the use of data mining process, information retrieval, technology-mediated learning environment, and visualization. Therefore, the definition of LA could be narrowed down to the statement that LA embraces certain features that can be directly projected into the research field with the use of computational techniques considering conditions and factors of the learning experience.

There are some review articles that summarize findings from multiple LA studies. These reviewed studies involved the use of LA as a research approach, which commonly involved the adoption of computational techniques. Papamitsiou and Economides (2014) systematically reviewed 209 published LA studies. Throughout the inclusion/exclusion procedure, they identified 40 key studies that involved the use of LA as a research approach. These researchers classified the studies by research strategy, research discipline, learning settings, research objectives, data collection technique, analysis technique, and results. They reported that the data in the reviewed LA studies were collected from different sources, such as log files, questionnaires, interviews, Google analytics, open datasets, and virtual machines. Specifically, they documented that LA researchers collected the data to measure the learner participation (e.g., login frequency, number of messages, forum and discussion posts), response times, task submission, previous grades in courses, detailed profiles, preferences, and affect observations (e.g., bored, frustrated, confused, happy). Also, they reported that the most frequently used methods were classification, followed by clustering, regression, and discovery with models. The algorithmic criteria computed for comparison of methods were precision, accuracy, sensitivity, coherence, fitness measures, and similarity weights. According to Papamitsiou and Economides's (2014) review article, which sheds light on the technical aspects, LA can be considered as a research approach that utilizes computational techniques for analyzing data collected from computer-based systems in order to investigate learner profile, behavior, and performance in the learning environment. Still, the use of computational techniques in research is not new. Thus, one of the key issues is the fact that there has not been a long enough period of time for educational researchers to utilize computational techniques. Then, what brings computational techniques to educational research?

Computational Techniques in Educational Research

From the review of previous LA studies wherein LA was used as a research method, it was found that LA studies process and/or visualize the data collected from interaction and navigation through computerized educational environments. LA studies apparently gained momentum from the availability of diverse data sources. The rapid development of communication and Internet technology has significantly changed how learning experiences are conceived and deployed. Numerous university programs today consist of online classes that include blended learning courses. The growing quantity of *big data* collected from the online learning environments during the past decade could not be handled manually. With the increased availability of large datasets and powerful computational engines, educational researchers began utilizing learners' behavior and experience to create insightful and real-time prediction models of learning processes. A few LA researchers acknowledge the benefits of the qualitative research approach because this approach provides rich descriptions of learning processes and additional information (Berland, Martin, Benton, Smith, & Davis, 2013; Chatti, Dyckhoff, Schroeder, & Thüs, 2012). However, few researchers have analyzed qualitative aspects in LA research.

LA researchers mostly collect students' activity data generated from technology-mediated learning systems, such as the number of clicks, discussion forums, assignments, test/assessments, and page views. Using these educational data, usually from learning management systems and other online learning platforms and tools (e.g., automatic assessment tools or intelligent tutoring systems) including MOOCs (Massive Online Open Courses), LA researchers aim to understand the learning process and to improve the quality of a learning experience in the learning systems (Pardo & Siemens, 2014). The data relating to learner experience and behavior also are supplemented with background information or profile of the learner. A few LA studies (e.g., Tempelaar, Rienties, & Giesbers, 2015) have involved the use of intentionally collected data, such as selfreport survey, along with the system-generated data; however, for most studies, LA can be seen as a case of the increased attention of the big data research phenomenon, and the utilization of computational analysis techniques (Ruipérez-Valiente, Muñoz-Merino, Leony, & Kloos, 2015).

The use of computational techniques for analyzing data collected from learning environments is one of the more important aspects when characterizing the concept of LA. Computational techniques have already been utilized in other areas, such as business intelligence, data mining, web analytics, and recommender systems. These fields include the investigation of big data handling techniques that can be used to analyze computer-readable sets of data (Persico & Pozzi, 2015; Serrano-Laguna, Torrente, Moreno-Ger, & Fernández-Manjón, 2014). The only difference between LA and these fields is that LA emerges as a link between the large data of learner experience in education settings and the computational techniques for analyzing learning-related data. Still, there are two similar areas (but two different names) under development that are oriented towards the inclusion and exploration of big data analysis within education: LA and Educational Data Mining (EDM).

EDM research could be traced back to the history of computer system development; yet, it is only in the late 2000 that EDM was recognized as a research field with attention aimed at how to utilize computer power in systematic ways for data analysis. EDM handles "developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would be hard or impossible to analyze due to the enormous volume of data within which they exist" (Romero & Ventura, 2013, p. 12). Apparently, this statement does sound like LA. Papamitsiou and Economides (2014) summarize the similarities of LA and EDM, as follows: (a) LA and EDM researchers collect, process, analyze, and report computer-readable data to advance learning/instructional process and the educational setting; (b) the research procedures of LA and EDM focus on the data collected from learning/instructional-related systems and preparation for processing during the instructors' and learners' activities; and (c) LA and EDM researchers analyze, report, and interpret the results in order to inform stakeholders (e.g., learners, instructors, organizations, and institutions) about learners' performance and the instructional/learning goal achievement, and, ultimately, to advise on the decision-making process of stakeholders (Papamitsiou & Economides, 2014). LA and EDM share similar goals and analysis methods/techniques aiming at investigating learning processes. Thus, some researchers do not differentiate between the two concepts (e.g., Berland et al., 2013; Mirriahi, Liaqat, Dawson, & Gašević, 2016). On the other hand, there is a perspective that both concepts are different from each other. For example, LA has been built upon a holistic viewpoint that focuses on understanding learning/instructional systems to their full complexity, whereas EDM involves a reductionistic stance that emphasizes analyzing individual components and new patterns in data, and modifying respective algorithms (Papamitsiou & Economides, 2014). However, as pointed out earlier, the holistic stance would define LA broadly; this, in turn, might not depict the unique characteristics of LA. In addition, from a methodological viewpoint, because the LA research area is sophisticated, the difference is becoming diluted. More importantly, the purposes of LA and EDM are especially similar to each other. Thus, it seems that there is no practical benefit in separating the two terms. Both of them involve employment of computational analysis techniques to understand learning and learning/instructional environments.

Purposes of Learning Analytics Studies

LA researchers aim to enhance the learning processes through systematic analysis of learning/instruction-related data and ultimately to provide informative feedback to stakeholders. In particular, LA provides stakeholders (e.g., learners, instructors, institutions) with opportunities to enable personalized learning. In many cases, for these purposes, LA researchers attempt to predict students' learning performances. For example, one purpose of typical LA studies is to identify or to predict learners who are (or will be) encountering obstacles in their learning, sometimes in real time (Serrano-Laguna et al., 2014). Predictions of dropout and retention are primary matters for LA researchers. In their review, Tempelaar et al. (2015) claim that a vast body of LA studies on student retention revealed that prediction models of the LA studies have well predicted students' academic performance via a range of demographic, academic integration, social integration, and psycho-emotional/social factors. In addition, researchers used the LA approach to monitor student interactions and individual assessment in diverse contexts (Fidalgo-Blanco, Sein-Echaluce, García-Peñalvo, & Conde, 2015). In some cases, visualization of the learners' online behaviors assists in improving both the teaching process and students' performance (Hernández-García, González-González, Jiménez-Zarco, & Chaparro-Peláez, 2015).

Research questions. When addressing the purpose of enhancing the learning processes, the LA approach has been used to answer particular research questions (see Table 1). Research questions about learners' profiles are one of the most frequently addressed questions (Mirriahi et al., 2016). Another type of research question deals with the changing of learners' behavior over time (Berland et al., 2013). The other type of research questions relates to prediction modeling (Nistor et al., 2014; Tempelaar et al., 2015). Currently, research relating to network analysis questions are frequently addressed in LA studies (Hernández-García et al., 2015). In addition, recent research has explored tools or mechanisms that can be integrated into the instructional system (Serrano-Laguna et al., 2014; Van Leeuwen, Janssen, Erkens, & Brekelmans, 2014).

Table 1. Research Questions in Learning Analytics Studies

Question types	Research	Questions
Learner profile	Mirriahi et al. (2016, p. 1088).	• What are the main learning profiles that emerge from the use of video annotation software? • Do different instructional methods influence the development of the learning profiles identified based on student engagement or use of video annotation software? • What is the effect of the learning profiles that emerge from the use of video annotation software on students' academic achievement?
Changes in learner behavior over time	Berland et al. (2013, p. 8).	• How, in the aggregate, does students' programming activity change over time? • What does this activity reveal about tinkering processes? • How do these changes relate to the quality of the programs that students are writing?
Prediction modeling	Tempelaar et al. (2015, p. 159)	• To what extent do (self-reported) learning dispositions of students, Learning Management Systems (LMSs), and e-tutorial data (formative assessments) predict academic performance over time? • To what extent do predictions based on these alternative data sources refer to unique facets of performance, and to what extent do these predictions overlap? • Which source(s) of data (learning dispositions, LMS data, e-tutorials formative tests) provide the most potential to provide timely feedback for students?
	Nistor et al. (2014, p. 340)	• (acceptance model verification) To what extent do acceptance factors (technology use intention, performance expectancy, effort expectancy, social influence, facilitating conditions and technology anxiety) predict participation in virtual community of practice? • (model verification) Does participation in virtual community of practice significantly mediate the influence of expertise on the expert status?
Network analysis	Hernández-García et al. (2015, p. 69)	• Are social network parameters of the different actors related to student outcomes in online learning? • Are global social network parameters related to overall class performance? • Can visualizations from social network analysis provide additional information about visible and invisible interactions in online classrooms that help to improve the learning process?
Tools or mechanisms	Serrano-Laguna et al. (2014, p. 317)	• Checking whether it was technically feasible to add the tracking mechanisms to a game that was developed separately • Testing whether a simple analysis of low-level interactions could be sufficient to identify game design issues and points in which the users were getting lost
	Van Leeuwen et al. (2014, p. 30)	• What is the effect of supporting tools that show information on student participation and discussion during collaboration on the development of teachers' diagnosis and interventions?

Overall, the various types of research questions in LA studies are somewhat limited. Given the limited data sources of a large amount of quantitative datasets, and the complex analysis techniques, the LA research area might have a limited boundary. However, the use of different technology-based learning environments and user-friendly analysis tools are increasing. Therefore, it is expected by the author that the LA approach will address a wide variety of research questions in the near future.

Research Process

The process of LA studies might vary, but there are some common steps and phases associated with these studies. A study conducted by Berland et al. (2013) can be used as an example. Specifically, this study exemplifies a robust and well-structured LA research process. Berland et al. (2013) used LA to understand how students learn computer programming through creative processes with computation. The data were collected from 53 female high school students learning to program (i.e., a game development summer camp) using a computer coding environment, which stores learners' behaviors during their game programming process. The primary data source was each state of the programs that the participants created. The measures were action (i.e., "the number of action primitives in a program state"), logic (i.e., "the total number of login and sensor primitives in a program state"), unique primitives (i.e., "the number of unique action, login, and sensor primitives in a program state"), length (i.e., "the total number of primitives in a program state"), coverage (i.e., "the percentage of possible combinations of sensor inputs"), and program quality (i.e., "how likely a student's robot is to win a game") (Berland et al., 2013, pp. 15-17). After confirming the measures, the researchers used a LA tool to conduct feature selection, which isolates particular features for inclusion in the analysis. Then, the participants' game program states were grouped into statistically generated categories (i.e., feature clustering). After identifying six different clusters, Berland et al. (2013) determined whether there were common sequences of a program state moved from one cluster to another over time. Then, the researchers interpreted the results within a current learning theory framework on computer coding (Berland et al., 2013).

As described earlier, there are specific analysis phases in LA studies. Pardo (2014) identified the following five independent steps: (a) Capture: the measurement, data collection (which might not directly ready to be processed), and LA technique selection; (b) Report: the data process using an arbitrary method ranging from simple to complex algorithms; (c) Predict: the answering stage for the previously formulated research questions; (d) Act: the generation of action that may change the target learning environment; and (e) Refine: the review of previous phases and adjustments to improve the suitability of each phase. By adopting these steps, the author of this article added an interpretation phase due to the author's aim of understanding LA as a research method in general (see Table 2).

For the present article, a literature search was conducted to identify the studies that used LA as a research approach. The ERIC Databases and Google Scholar were queried to search for literature in this field wherein LA studies were published in the area of educational research. Titles, abstracts, and keywords were searched for Learning Analytics. The initial set of 131 references narrowed via the following first selection criterion: the study had to involve collection and analysis of data with a LA technique(s). The author of this study identified 64 articles wherein LA was used as a research approach for the actual analysis. The literature search identified 14 studies that were obtained for a full review after the second screening criterion: the study provides clear descriptions of their analysis process. As can be seen in Table 2, the studies utilized different types of methods in each phase (i.e., data collection, feature selection, technique application, and interpretation). Interestingly, the studies that involved data collected from a similar type of data source adopted similar approaches in other phases (i.e., feature selection, technique application, and interpretation). For example, the studies that involved the use of a learning management system as a data source involved a selection of environment-specific features and then involved the use of clustering techniques for exploratory purposes.

Overall, all data sources were computerized learning environments. Although there existed some variations on the use of strategies in each phase, the reviewed studies involved an adoption of computational techniques for the analysis following the four steps: data collection, feature selection, technique application, and interpretation.

Methodological Issues in Learning Analytics Research

The rapid development of technology supports the collection of vast amounts of data and their resulting analysis and reporting, which brings computational techniques into the educational research method field. However, there are methodological issues to address before LA becomes a robust research approach. In the research methodology field, Mertens et al. (2016) have already raised issues of the trend (i.e., computational analysis of the large amount of data) within a mixed methods perspective by denoting some concerns, such as the quality of big data, its relevance, feature selection, data preprocessing (e.g., data merge), levels of analysis, and confidentiality. Specifically, these researchers designated the issue of how to integrate the computational results with a qualitative component. In addition, due to big data being considered as a population of a certain context, mixed methods researchers might have to think about the reconceptualization of sample, sample size, and entire population (Mertens et al., 2016). These concerns also apply to the LA approach. For example, LA faces the issue of qualitative data integration (Chatti et al., 2012) or confidentiality (Pardo & Siemens, 2014). In addition to those issues listed, more methodological thoughts on the LA approach are addressed in the following section.

Table 2. Learning Analytics Study Phases

Phase	Requirement	Application
Data collection	Measure determination and definition	• Programming software (Berland et al., 2013; Blikstein, 2011) • Online community of practice (Nistor et al., 2014) • Video annotation software (Mirriahi et al., 2016) • Online concept mapping tool (Scheffel et al., 2014) • Educational video game (Serrano-Laguna et al., 2014) • Learning management system (Agudo-Peregrina, Iglesias-Pradas, Conde-González, & Hernández-García, 2014; Fidalgo-Blanco et al., 2015; Hernández-García et al., 2015; Lust, Elen, & Clarebout, 2013; Lust, Vandewaetere, Ceulemans, Elen, & Clarebout, 2011; Tempelaar et al., 2015) • Virtual math software (Xing, Guo, Petakovic, & Goggins, 2015) • Online discussion forum (Wise, Zhao, & Hausknecht, 2013)
Feature selection	Selection justification	• Data collection environment-specific (Agudo-Peregrina et al., 2014; Berland et al., 2013; Blikstein, 2011; Lust et al., 2013; Lust et al., 2011; Mirriahi et al., 2016; Serrano-Laguna et al., 2014; Wise et al., 2013) • Research topic-specific (Fidalgo-Blanco et al., 2015; Hernández-García et al., 2015; Xing et al., 2015) • Literature review (Nistor et al., 2014; Tempelaar et al., 2015) • Expert review (Scheffel et al., 2014)
Technique application	Alignment with research questions	• Clustering (Agudo-Peregrina et al., 2014; Berland et al., 2013; Lust et al., 2013; Lust et al., 2011; Mirriahi et al., 2016; Scheffel et al., 2014) • Regression (Agudo-Peregrina et al., 2014; Nistor et al., 2014; Tempelaar et al., 2015) • Path analysis (Berland et al., 2013) • Tracking and behavior analysis (Blikstein, 2011; Fidalgo-Blanco et al., 2015; Serrano-Laguna et al., 2014; Wise et al., 2013) • Social network analysis and visualization (Hernández-García et al., 2015) • Genetic algorithm (Xing et al., 2015)
Interpretation	Alignment with research context	• Framework formulation (Berland et al., 2013; Scheffel et al., 2014) • Prediction modeling (Hernández-García et al., 2015; Tempelaar et al., 2015; Xing et al., 2015) • Existing model validation (Fidalgo-Blanco et al., 2015; Nistor et al., 2014) • Instructional design assessment (Serrano-Laguna et al., 2014) • Exploratory analysis (Agudo-Peregrina et al., 2014; Blikstein, 2011; Lust et al., 2013; Lust et al., 2011; Mirriahi et al., 2016; Wise et al., 2013)

Interpretability

One of the primary purposes of LA studies is to formulate a learner performance prediction model. This prediction model must offer practical and realistic guidance to learners, instructors, curriculum developers, and administrators in order to improve learners' levels of performance and academic achievement. This raises an issue of result interpretation and data contextualization. Xing et al. (2015) conceptualized the interpretability issue as the black/white box of LA techniques. The main assumption of this argument is that LA studies need to be designed to provide meaningful and interpretable prediction models that are easily understandable at the practitioner level, and which do not require any types of sophisticated knowledge about computational techniques in order to use the suggested model. According to the researchers, there are two types of LA methods: white-box and black-box: White-box methods are easily understood and interpreted by persons who do not have any specific background knowledge about computer programming or statistics, as opposed to black-box methods, which are difficult (almost impossible) to be comprehended by practitioners who do not have the background knowledge about the analysis techniques (Xing et al., 2015). Even if LA studies suggested useful insights on the learning process and its mechanism, the information and results provided to the instructors would not always be straightforward to interpret. Thus, it is recommended that LA researchers attempt to provide specific implications and practical guidance for an audience of laypersons even though black-box methods were used in their research.

Methodological Advances

Modeling students' learning process in educational settings is not new. Traditional modeling techniques, including linear or logistic regression, have been successfully utilized for a variety of educational studies. Instructional/Learning environments have evolved, such as learning management systems, MOOCs, online courses, and complex artifacts. As they require deeper analysis, the traditional modeling techniques have shown some limitations. Specifically, these limitations include a lack of established paradigm for optimizing learning performance prediction (further explained in Xing et al., 2015). To fill the gap, LA has been suggested. Notwithstanding, LA is neither self-explanatory nor self-regulatory. Although research approaches, methods, and techniques might lead the researcher to a certain stance of learning and assessment, the data sources should not determine the research direction. Thus, researchers need to be informed of the various ways that LA can be adopted, depending on the researcher's philosophical standpoint in different pedagogical contexts. For this, five philosophical stances with appropriate LA approaches offered by Knight et al. (2014) would be beneficial: (a) Constructivism (in its various forms): The focus of LA is on learners' progress through tracking learners' behaviors and making decision on the instructional modification, such as instructional materials, resources, and tools; (b) Subjectivism: The focus of LA is on learners' motivational aspects, such as understanding why a learner is (or is not) engaged in a particular learning task; (c) Apprenticeship: The focus of LA is on the classification of expert and novice learners with underlying reasons, and the knowledge transfer or shift between them; (d) Connectivism: The focus of LA is on the network analysis (e.g., networks' size, node, quality, changes over time) in order to investigate the connection of concepts and knowledge; and (e) Pragmatism (in its various forms): The focus of LA is on the learning process rather than on the learning performance (Knight et al., 2014). Following these guidelines might restrict the researcher's creativity. However, as a methodology, the LA approach must also play a significant role in facilitating educational researchers in locating unknown patterns within data rather than "proceeding from a query initiated by a traditionally testable hypothesis" (Berland et al., 2013, p. 8).

Researchers of recent LA publications advocate for additional discussion about the soundness and suitability of their analysis techniques. This effort should lead to more practical guidance regarding how each LA technique can be fully utilized. Although there are similarities among different analysis techniques in LA—for example, classification and clustering include similar algorithms—each technique offers a unique set of advantages and disadvantages (e.g., the use of supervised—labeled data in the classification technique, or the use of unsupervised—unlabeled data in clustering technique) (Martin & Sherin, 2013). Emphasizing how LA techniques can be used in a certain condition would then guide future researchers to elaborate on how to utilize a technique in a specific educational context, which would play a significant role in enhancing LA as a robust research method. In addition, feature selection methods should be further investigated (see the second phase of LA Study Phases in Table 2). When selecting features among many types of indicators, researchers in this area normally review the literature to identify features that were selected in previous research. Although the literature review method is convenient, it should be noted that the feature selection process is highly context-specific (Ruipérez-Valiente et al., 2015). Thus, researchers should be able to consider their own research context first, then, to identify appropriate feature selection methods and/or algorithms. Due to few studies suggesting an enhanced method for feature selection, more effective feature selection approaches also need to be further investigated, and the practical and detailed guidance for feature selection is required to enhance LA as a robust research method.

Challenges or Recommendations

In this article, challenges of LA studies have been identified. For the LA approach to become a robust method for educational research, the following challenges need to be overcome.

Connecting with educational theory. In order to gain a deeper understanding of the features that impact learning and performance, researchers should be able to contextualize their data using educational theories or principles (Xing et al., 2015). Because a large set of variables might diminish statistical prediction power, the feature selection is a critical phase in LA. As opposed to the computer science or mathematical field that employs computational methods for selecting features, the researcher's judgment is the key point of feature selection in educational research. The decision-making process can be aided by educational theories or principles. In addition, LA researchers should be able to identify a relationship between the study results and the educational theories/principles rather than to interpret the results as they are. This is because the results of educational research are meaningful only when they accumulate knowledge of actual instruction/learning design (Tempelaar et al., 2015). However, it is strenuous to identify an appropriate educational theory or principle for both feature selection and result interpretation. Thus, LA researchers should scrutinize pedagogical grounds prior to the research being conducted.

Context and content knowledge. In addition to the pedagogical framework based on educational theories, our knowledge from LA can only be considered evidence in a contextualized instance, such as in a specific online learning environment. The learning environment as a context should be "redefined as a metacognitive tool which cannot be excluded in assessment" (Knight et al., 2014, p. 6). Without understanding the learning environment as well as its functionality, accessibility, usability, and other technical aspects, it would not be possible to grasp the core of learning experience and to distribute any meaningful implications of the LA study results. Thus, more contextual knowledge about the learning environment is required for the LA approach, and the detailed information about the context should be described in the LA study. In addition, in order for the researcher to utilize the LA approach, computational literacy, including proper use of computational tools and knowledge about the computational techniques, is necessary.

Visualization and broader perspective. Ultimately, LA researchers aim to inform decision makers about different stages of events in the learning process across educational institutions. In this way, LA contributes to the educational research field by providing empirical support for previously theorized processes and presenting data-driven evidence. LA informs decision-making processes regarding curriculum and instructional design, potentially providing different useful solutions when compared to traditional research approaches in education (Berland et al., 2013). The informed decision also might involve the revision and fine-tuning of the instructional system under development (Persico & Pozzi, 2015). As per the earlier discussion regarding the interpretability issue, more practical and comprehensible implications must be delivered to the stakeholders and decision makers. In many cases, providing stakeholders with visual information can better support their decision-making process. Thus, selecting suitable and effective graphical data representation methods is another matter for consideration for future LA research (Persico & Pozzi, 2015).

Ethical issues. Because the data sources for LA studies are diversified, and the data types are innumerable, researchers might overlook some ethical issues. Pardo and Siemens (2014) pointed out several ethical issues of LA studies, such as personal background information, sharing delicate information, learner privacy, possibly identifiable information in the publication, and so forth. Then, they contemplated possible guidelines for the ethical issues: transparency of every stage in the LA process, learner control over data, defined right of access to all the collected data, and accountability or robustness of the overall process (Pardo & Siemens, 2014, pp. 445-448). In the same vein, LA researchers should be cautious when processing the data, particularly factors that are related to the students' personal information.

Conclusion

The field of LA has contributed to the advancement of understanding regarding the technology-oriented learning environments as well as guiding instructors, learners, and researchers for an enhanced learning performance. However, knowledge is still lacking about the topological status of LA as a research approach. Few researchers have discussed methodological issues when addressing LA as a research approach. Methodologically speaking, the LA approach must forecast learner behavior and performance by focusing on the appropriate techniques, algorithms, and methods, along with deep consideration of educational context, theories, and phenomena. Among the challenges facing the LA approach is the need for the appropriate use of techniques and how they are interpreted in real educational settings. The challenges associated with LA studies also include how to incorporate findings from a prediction model into current educational theories or principles, as well as how to tie the findings into the educator's decision-making process regarding more practical manners. Thus, further questions for a robust research approach that the LA researchers should consider include: What is the level of understanding of LA and what is the usefulness of each LA technique for furthering understanding of methodology? Additional questions about advancing the LA approach include how can LA researchers integrate a variety of techniques at all levels of the research study—that is, at the philosophical and theoretical level, for data collection and analysis, and for reporting and use?

Finally, while enhancing the results' capacity to be responsive to stakeholders for better decision-making process, the LA approach requires more flexibility and creativity, which might be acquired by incorporating qualitative research approaches. Within the analysis process, with mixed methods research approaches, it can also be argued that encouragement of creativity and openness to new ideas are necessary for the LA approach to advance. In this way, it is expected that the LA approach will be enhanced by the development of new methodologies and approaches to address increasingly complex educational research questions. As learning is becoming more individualized and the learning context is diversified, educational problems can be considered as *wicked problems*. Wicked problems refer to "problems involving multiple interacting systems, replete with social and institutional uncertainties, for which there is no certainty about their nature and solutions, and for which time is running out to find solutions" (Mertens et al., 2016, p. 225). Wicked problems would be addressed by utilizing mixed methods research (Chestnut, Hitchcock, & Onwuegbuzie, in press), if mixed methods research approaches involved innovations in methodology (Mertens et al., 2016). Therefore, the LA community should consider actively incorporating qualitative methods into the LA approach, which would facilitate LA researchers in addressing wicked educational problems.

References

1st International Conference on Learning Analytics and Knowledge 2011. (2010). *About*. Retrieved from https://tekri.athabascau.ca/analytics/about

Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior, 31,* 542-550. doi:10.1016/j.chb.2013.05.031

Berland, M., Martin, T., Benton, T., Smith, C. P., & Davis, D. (2013). Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences, 22,* 564-599. doi:10.1080/10508406.2013.836655

Blikstein, P. (2011). Using learning analytics to assess students' behavior in open-ended programming tasks. In P. Long, G. Siemens, G. Conole, & D. Gašević (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 110-116). New York, NY: ACM.

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

Chestnut, C., Hitchcock, J. H., & Onwuegbuzie, A. J. (in press). Using mixed methods to inform education leadership and policy research. In C. R. Lochmiller (Ed.), *Complementary research methods in educational leadership and policy studies*. London, England: Palgrave Macmillan.

Fidalgo-Blanco, Á., Sein-Echaluce, M. L., García-Peñalvo, F. J., & Conde, M. Á. (2015). Using learning analytics to improve teamwork assessment. *Computers in Human Behavior, 47,* 149-156. doi:10.1016/j.chb.2014.11.050

Hernández-García, Á., González-González, I., Jiménez-Zarco, A. I., & Chaparro-Peláez, J. (2015). Applying social learning analytics to message boards in online distance learning: A case study. *Computers in Human Behavior, 47,* 68-80. doi:10.1016/j.chb.2014.10.038

Knight, S., Shum, S. B., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47. doi:10.18608/jla.2014.12.3

Lust, G., Elen, J., & Clarebout, G. (2013). Regulation of tool-use within a blended course: Student differences and performance effects. *Computers and Education, 60*(1), 385-395. doi:10.1016/j.compedu.2012.09.001

Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., & Clarebout, G. (2011). Tool-use in a blended undergraduate course: In search of user profiles. *Computers and Education, 57,* 2135-2144. doi:10.1016/j.compedu.2011.05.010

Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences, 22,* 511-520. doi:10.1080/10508406.2013.840466

Mertens, D. M., Bazeley, P., Bowleg, L., Fielding, N., Maxwell, J., Molina-Azorin, J. F., & Niglas, K. (2016). Expanding thinking through a kaleidoscopic look into the future: Implications of the Mixed Methods International Research Association's Task Force report on the future of mixed methods. *Journal of Mixed Methods Research, 10,* 221-227. doi:10.1177/1558689816649719

Mirriahi, N., Liaqat, D., Dawson, S., & Gašević, D. (2016). Uncovering student learning profiles with a video annotation tool: Reflective learning with and without instructional norms. *Educational Technology Research and Development, 64,* 1083-1106. doi:10.1007/s11423-016-9449-2

Nistor, N., Baltes, B., Dascălu, M., Mihăilă, D., Smeaton, G., & Trăuşan-Matu, Ş. (2014). Participation in virtual academic communities of practice under the influence of technology acceptance and community factors. A learning analytics application. *Computers in Human Behavior, 34,* 339-344. doi:10.1016/j.chb.2013.10.051

Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology and Society, 17*(4), 49-64.

Pardo, A. (2014). Designing learning analytics experiences. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 15–38). New York, NY: Springer.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technol-*

ogy, 45, 438-450. doi:10.1111/bjet.12152

Persico, D., & Pozzi, F. (2015). Informing learning design with learning analytics to improve teacher inquiry. *British Journal of Educational Technology, 46,* 230-248. doi:10.1111/bjet.12207

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3*(1), 12-27. doi:10.1002/widm.1075

Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Kloos, C. D. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior, 47,* 139-148. doi:10.1016/j.chb.2014.07.002

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Educational Technology and Society, 17*(4), 117-132.

Serrano-Laguna, Á., Torrente, J., Moreno-Ger, P., & Fernández-Manjón, B. (2014). Application of learning analytics in educational videogames. *Entertainment Computing, 5,* 313-322. doi:10.1016/j.entcom.2014.02.003

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior, 47,* 157-167. doi:10.1016/j.chb.2014.05.038

Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education, 79,* 28-39. doi:10.1016/j.compedu.2014.07.007

Wise, A. F., Zhao, Y., & Hausknecht, S. N. (2013). Learning analytics for online discussions: A pedagogical model for intervention with embedded and extracted analytics. In D. Suthers & K. Verbert (Eds.), *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 48-56). New York, NY: ACM Press. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.824.454&rep=rep1&type=pdf

Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable genetic programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47,* 168-181. doi:10.1016/j.chb.2014.09.034