A Classification Framework for Research on Learning Analytics and a Literature Review with a Focus on Professional Learning

Kibong Song
Center for Instructional Technology Solutions in Industry and Education
Virginia Polytechnic Institute and State University
144G Smyth Hall (0488), 185 Ag-Quad Ln, Blacksburg, VA 24061, USA
gibbs@vt.edu

Kenneth Potter
Instructional Design and Technology
Virginia Polytechnic Institute and State University
144J Smyth Hall (0488), 185 Ag-Quad Ln, Blacksburg, VA 24061, USA
kpotter@vt.edu

Barbara Lockee
Instructional Design and Technology
Virginia Polytechnic Institute and State University
330 Burruss Hall (0132), 800 Drillfield Dr, Blacksburg, VA 24061, USA
lockeebb@vt.edu

Descriptors: Learning Analytics, Classification Framework

Abstract

Many researchers have studied learning analytics during the last decade. However, obstacles still exist that impede enhancement of the field. These include a conceptual misunderstanding of learning analytics, an imbalance of research scope, and an inclusion of unrelated research. In this study, we suggest a classification framework for research on learning analytics. We analyzed 608 articles from the literature and found patterns regarding the definitions of learning analytics, contexts of the studies, methodologies utilized for the studies, and scopes of the studies. Based on the patterns, we developed a classification framework, consisting of four layers: Foundation Layer, Environment Layer, Development Layer, and Application Layer. We classified articles on learning analytics for professional learning and identified research gaps. The results show that the research on learning analytics for professional learning is not balanced in terms of scope. The conclusions and limitations are discussed.

Introduction

Analytics as the science of analysis is not a new concept. Even before anno domini, people analyzed patterns of clouds and winds to predict weather. However, modern analytics can be interpreted differently. With the development of computing systems and advancement of statistics, analytics is based mostly on proven statistical models and data collected by computing systems. Given these technological and technical changes, it is generally taken for granted that modern analytics refers to "the process of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data" (Cooper, 2012, p.3). Modern analytics began in the 1980s. Though some companies started to use computing systems to manage customer data in the 1970s, database marketing and customer relationship management software in the 1980s were the beginning of modern analytics.

Technologies, such as the Internet, to increase accessibility of various people and reduction in costs associated with using analytics contributed to the expansion of the usage of analytics to educational contexts and made many researchers and practitioners in the field interested in analytics for education. To reflect such interests, researchers have introduced academic analytics (Campbell, DeBlois, & Oblinger, 2007; Goldstein, 2005; van Barneveld, Arnold, & Campbell, 2012) and learning analytics (Ferguson, 2012; Long, Siemens, Conole, & Gašević, 2011). While academic analytics supports decision-making in regard to an academic organization and its performance, learning analytics more focuses on an individual student's learning and performance.

Learning analytics (LA) is analytics for supporting decision-making regarding learning and learning environments. During the past decade, researchers from various disciplines have studied LA. While the diversity of disciplines has the advantage of expanding the research and introducing different viewpoints, it can be an obstacle to interdisciplinary collaboration if the interpretation of a particular concept varies from discipline to discipline. For example, researchers in one field regard LA as an automated tool that supports learning, whereas researchers in another field see LA as a process to analyze data on learning and learning environments. Another problem is that there are imbalances. For example, researchers have paid less attention to the LA for professional learning. At this point, it seems meaningful to try to find patterns and gaps in LA research to overcome those problems.

This study has three purposes. The first is to find patterns of research on LA in the following respects: 1) how researchers have defined LA, 2) in what contexts researchers have conducted their research, 3) what methodologies researchers have used, and 4) what the research scopes are. The second is to develop a classification framework for LA research based on the patterns. The last purpose is to identify gaps in research on LA for professional learning. The following section introduces a literature review on LA and classification frameworks for LA research.

Literature Review

In this section, we provide a review of the literature on classification frameworks for LA or LA research, consisting of three viewpoints. We discuss studies on the classification frameworks for LA or LA research based on these viewpoints. Later in this section, we introduce the research questions of this study.

Environment Viewpoint

The environment viewpoint focuses on environmental factors and conditions related to LA. Some researchers (Greller & Drachsler, 2012; Peña-Ayala, 2018) investigated LA or LA research from the environment viewpoint. Peña-Ayala (2018) proposed a classification framework for LA research, for which the researcher explained three key areas: profile, applications, and underlying factors. In the framework, the researcher viewed legal issues, theoretical topics, and learning paradigms and settings as underlying factors influencing LA. The researcher also classified the definition of LA, stakeholders, field evolution, underlying domains, related domains, specialized lines, and prior reviews of the LA field as profile that reveals "an overall perspective of what LA is" (Peña-Ayala, 2018, p. 4). Greller and Drachsler (2012) also considered the environment factors and conditions in designing a framework for the domain of LA. Their framework consists of six dimensions: stakeholders, internal limitations, external constraints, instruments, data, and objectives. The first five dimensions are associated with environments surrounding LA while the objective is closely related to the application of LA. The frameworks of Greller and Drachsler (2012) and Peña-Ayala (2018) well reflect theoretical fundamentals of LA research as well as environmental factors and conditions related to LA.

Development Viewpoint

The development viewpoint focuses on elements that are necessary for developing LA. Aljohani et al. (2019), Muslim, Chatti, Bashir, Varela, and Schroeder (2018), Yassine, Kadry, and Sicilia (2016, April), and Ifenthaler and Widanapathirana (2014) studied frameworks for LA and we categorized them as the framework reflecting the development viewpoint. While the framework by Ifenthaler and Widanapathirana (2014) covers a broader scope and more elements of LA, the others emphasize the core functions of LA.

Aljohani et al. (2019) proposed a course-adapted student learning analytics framework, which consists of four levels: instructor, data, data analytics, and presentation levels. This framework focuses on the analytical process from collecting data on learning and learning environments through presenting information such as feedback. Muslim et al. (2018) also focused on the analytical process in developing a modular framework for open learning analytics, but their framework includes more detailed processes and interactions among modules of the analytical process. The modular framework consists of four modules: analytics engine, analytics modules, analytics methods, and visualizer. Yassine, Kadry, and Sicilia (2016, April) considered user activities and learning outcomes in addition to data analysis and visualization in their framework. Their framework contains the definitions of data on user activities, mapping activities with learning outcomes, analysis of data on activities and learning outcomes, and information visualization.

Unlike the aforementioned researchers, Ifenthaler and Widanapathirana (2014) introduced a holistic framework for LA, in which three core engines (learning analytics engine, personalization and adaptation engine, and reporting engine), use various data generated by different sources (individual characteristics, social web, physical data, curriculum, and online learning environment), to provide information to the stakeholders including institution or governance. Though the holistic framework includes only institution and governance as stakeholders and the three engines in the framework are overlapped (e.g. visualization) and less relevant to the core features of LA (e.g. gamification), it has contributed to the field of LA in that it covers various data sources (e.g. social web and physical data) and separates functions of LA into learning analytics engine, personalization and adaptation engine, and reporting engine.

Application Viewpoint

The application viewpoint focuses on the uses and practical applications of LA. Many researchers in the field of LA studied the applications of LA and their studies highlighted specific purposes of LA, such as prediction of learning performance and retention (Hicks, 2018; Lu et al., 2018; Marbouti, Diefes-Dux, & Madhavan, 2016; Yu et al., 2018) or understanding of learners' behaviors (Berland, Martin, Benton, Smith, & Davis, 2013; Martín-Monje, Castrillo, & Mañana-Rodríguez, 2018; Ruipérez-Valiente, Muñoz-Merino, Leony, & Kloos, 2015), rather than discussing classification of applications of LA. There is not enough research on classification frameworks for LA seen from the application viewpoint. For this reason, we expanded our literature review to other data analytics areas.

Some researchers (Fleckenstein & Fellows, 2018; Kumar, 2017; Skourletopoulos, Mastorakis, Mavromoustakis, Dobre, & Pallis, 2018) argued that analytics can be classified as descriptive, diagnostic, predictive, and prescriptive analytics based on the types of analytics applications. Their studies classified the types of analytics applications based on what kinds of information analytics can provide. Descriptive analytics shows information describing what happened or what is happening. Diagnostic analytics explains causal relationships by reporting information on why it happened or what it is happening. Predictive analytics provides information on what might happen. Prescriptive analytics recommends interventions based on the prediction.

Rationale of the Research and Research Questions

While there have been a few articles on classification frameworks for LA, frameworks for LA research remain under-researched. To promote the advancement of research on LA, it is necessary to develop a classification framework for LA research. Given the purposes of the study mentioned in the previous section, we addressed the following research questions: 1) what patterns exist in the definitions of LA, 2) what patterns exist in the contexts of research on LA, 3) what patterns exist in the methodologies used for research on LA, 4) what patterns exist in the scopes of research on LA, 5) how can the classification framework for LA research be developed, and 6) what are some gaps in research on LA for professional learning? In order to answer these research questions, we employed the research methodology illustrated in the following section.

Research Methodology

This study is a mapping review, which is a type of study that seeks "to map out and categorize existing literature on a particular topic, identifying gaps in research literature from which to commission further reviews and/or primary research" (Grant & Booth, 2009, p. 97). For the mapping review, we searched multiple academic databases to find articles to be reviewed. In this section, we explain our research methodology by describing articles selection process, information sources and search strategy, and exclusion criteria.

Articles Selection Process

For this mapping review, we searched for articles from five academic databases: ERIC from EBSCOhost, PsycINFO, Academic Search Complete from EBSCOhost, Education Research Complete from EBSCOhost, and Web of Science from Clarivate Analytics. We searched these databases on August 1st, 2018 and found 1,467 articles in total. After removing 754 duplicates, 713 articles remained. By applying the exclusion criteria explained later in this section, 608 articles remained for review. *Figure 1* illustrates a summary of articles selection process for this study.

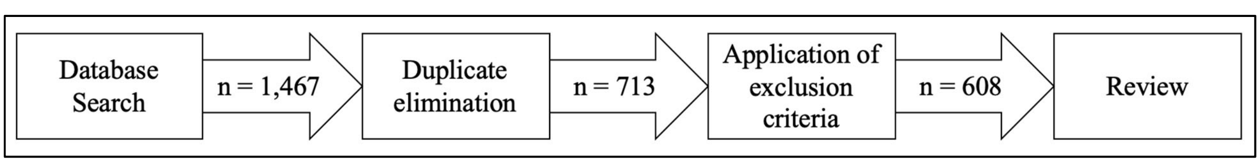

Figure 1. Summary of articles selection process

Information Sources and Search Strategy

As mentioned above, we used five academic databases as information sources for this study. Since the search feature of each database is slightly different, we used different search conditions for each database. Table 1 shows the respective search conditions for each database.

Table 1. Search conditions for each database

Database	Search conditions	No. of articles
ERIC from EBSCOhost	- Search keywords: 'learning analytics' for title or 'learning analytics' for KW identifiers - Only peer reviewed - Only journal articles for publication type - Find all my search terms	157
PsycINFO	- Search keywords: 'learning analytics' for title or 'learning analytics' for keywords - Only peer reviewed - Only journal articles for document type	167
Academic Search Complete from EBSCOhost	- Search keywords: 'learning analytics' for title or 'learning analytics' for KW Identifiers - Only peer reviewed - Only journal articles for document type - Find all my search terms	249
Education Research Complete from EBSCOhost	- Search keywords: 'learning analytics' for title or 'learning analytics' for KW Identifiers - Only peer reviewed - Only journal articles for document type - Find all my search terms	378
Web of Science from Clarivate Analytics	- Search keywords: 'learning analytics' for title (TI) or 'learning analytics' for topic (TS) - Only Articles for document type - Only from Web of Science Core Collection - Only from SCIE, SSCI, A&HCI, and ESCI (Emerging Sources Citation Index)	516

Exclusion Criteria

Given the purposes of the study, we adopted a comprehensive strategy in setting exclusion criteria. We did not include the quality of an article or a journal in the criteria. We excluded articles that are not written in English, not related to LA, or not research (e.g. editorial, commentary, or book review). Based on the exclusion criteria, we excluded 26 articles not written in English, 38 articles not related to LA, and 41 articles that are not research.

Results

Research Question 1: What patterns exist in the definitions of LA?

Among 608 articles, we analyzed 185 articles that defined LA or used existing definitions of LA and found 31 different definitions of LA. We included only six definitions in our analysis based on the frequency of the definition usage and excluded 25 definitions as they have been used only one time. Table 2 illustrates the definitions of LA that have been used in the articles we reviewed more than one time.

Table 2. Definitions of learning analytics and frequency

Author(s) and publication year	Definition	No. of usage
Society for Learning Analytics Research (2011)	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	n = 154 (77%)
Siemens (2010)	Use of learner-produced intelligent data and analysis models to discover information and social connections and to predict and advise on learning activities	n = 12 (6%)
Slade & Prinsloo (2013)	Collection, analysis, use, and appropriate dissemination of student-generated, actionable data with the purpose of creating appropriate cognitive, administrative, and effective support for learners	n = 3 (1.5%)
Lawson, Beer, Rossi, Moore, & Fleming (2016)	Collection and analysis of data in education settings in order to inform decision making and improve learning and teaching	n = 2 (1%)
Willis, Zilvinskis, & Borden (2017)	Process of using live data collected to predict student success, promote intervention or support based on those predictions, and monitor the influence of that action	n = 2 (1%)
Brown (2011)	Collection and analysis of usage data associated with student learning and its purpose is to observe and understand learning behaviors in order to enable appropriate interventions	n = 2 (1%)

The most used definition of LA is "the measurement, collection, analysis and reporting of data about learners and their contexts, for the purpose of understanding and optimizing learning and the environments in which it occurs" (Society for Learning Analytics Research, 2011). 77% of the used LA definitions referenced the definition of the Society for Learning Analytics Research. Siemens's definition of LA (2010) is the next most frequently used definition (6%), followed by Slade and Prinsloo's definition (1.5%), Lawson, Beer, Rossi, Moore, and Fleming's definition (1%), Willis, Zilvinskis, and Borden's definition (1%), and Brown's definition (1%).

These definitions share commonalities. All the definitions consist of three components: data type, process, and purpose of LA (see Table 3). First, the definitions, except for the definition by Willis et al. (2017), clearly indicate that the data for LA are associated with learners, their learning, or their contexts. Second, the definitions describe the process related to LA. The process includes collection, analysis, and reporting steps. While the definitions by Siemens (2010) and Willis et al. (2017) do not specify steps during the process, the other definitions indicate each step during the process in a sequential manner. Lastly, all the definitions include purposes of LA. While the definitions by Siemens (2010), Slade and Prinsloo (2013), Willis et al. (2017), and Brown (2011) include purposes that are directly associated with functions of LA, the definitions by the Society for Learning Analytics Research (2011) and Lawson et al. (2016) include indirect purposes, such as 'optimizing learning and the environments' and 'improving learning and teaching.'

Table 3. Data types, processes, and purposes in LA definitions

Author(s) and publication year	Data type	Process	Purpose
Society for Learning Analytics Research (2011)	Data about learners and their contexts	Measurement, collection, analysis and reporting	Understanding and optimizing learning and the environments in which it occurs
Siemens (2010)	Learner-produced intelligent data	Use of … data and analysis models	To discover information and social connections and to predict and advise on learning activities

Slade & Prinsloo (2013)	Student-generated, actionable data	Collection, analysis, use, and appropriate dissemination	Creating appropriate cognitive, administrative, and effective support for learners
Lawson et al. (2016)	Data in education settings	Collection and analysis	In order to inform decision making and improve learning and teaching
Willis et al. (2017)	Using live data	Process of …	To predict student success, promote intervention or support based on those predictions, and monitor the influence of that action
Brown (2011)	Data associated with student learning	Collection and analysis of	To observe and understand learning behaviors in order to enable appropriate interventions

Research Question 2: What patterns exist in the contexts of research on LA?

We found 612 contexts from 608 articles and classified the contexts into five categories: higher education, K-12, MOOC-based learning, professional learning, and others. The 'Others' category includes articles that did not clearly indicate a context for research. Table 4 presents the results of our classification. As shown in Table 4, the most frequently studied context for LA research is higher education, occupying 300 out of 612 contexts, followed by K-12 context (n=69), MOOC-based learning context (n=25), and professional learning context (n=11).

Table 4. Changes of contexts in LA research

Context\Year	2011	2012	2013	2014	2015	2016	2017	2018*	Total
HE	1	6	9	20	52	62	91	59	300
K-12	0	0	3	7	17	16	18	8	69
MC	0	0	0	0	4	9	8	4	25
PL	0	0	2	0	1	3	3	2	11
Others	0	1	3	2	7	4	5	3	25
Total	1	14	27	51	110	147	170	92	612

Note. HE = Higher education; PL = Professional learning; MC = MOOC-based learning. The figures for 2018 are the number of contexts found in the articles published between January and July 2018.

Figure 2 illustrates changes in the relative ratio of the contexts in LA research reviewed. The results revealed that there is a significant imbalance in research on LA in terms of research contexts. Many researchers have studied LA in higher education contexts, but relatively few researchers have studied LA in other contexts, notably professional learning. The ratios for the higher education context have always been or higher than 64.3%. On the other hand, the ratios for the professional learning context have been less than 3.4%, except for 2013 (14.3%).

In addition to the imbalance of contexts, the results discovered another pattern regarding the research context. Since 2015, researchers in the field have studied LA in the MOOC-based learning context. The relative ratios for the MOOC-based learning context in LA research have been 5.4% or higher. The last pattern found in the results is that the K-12 learning context has been steadily decreasing since 2014.

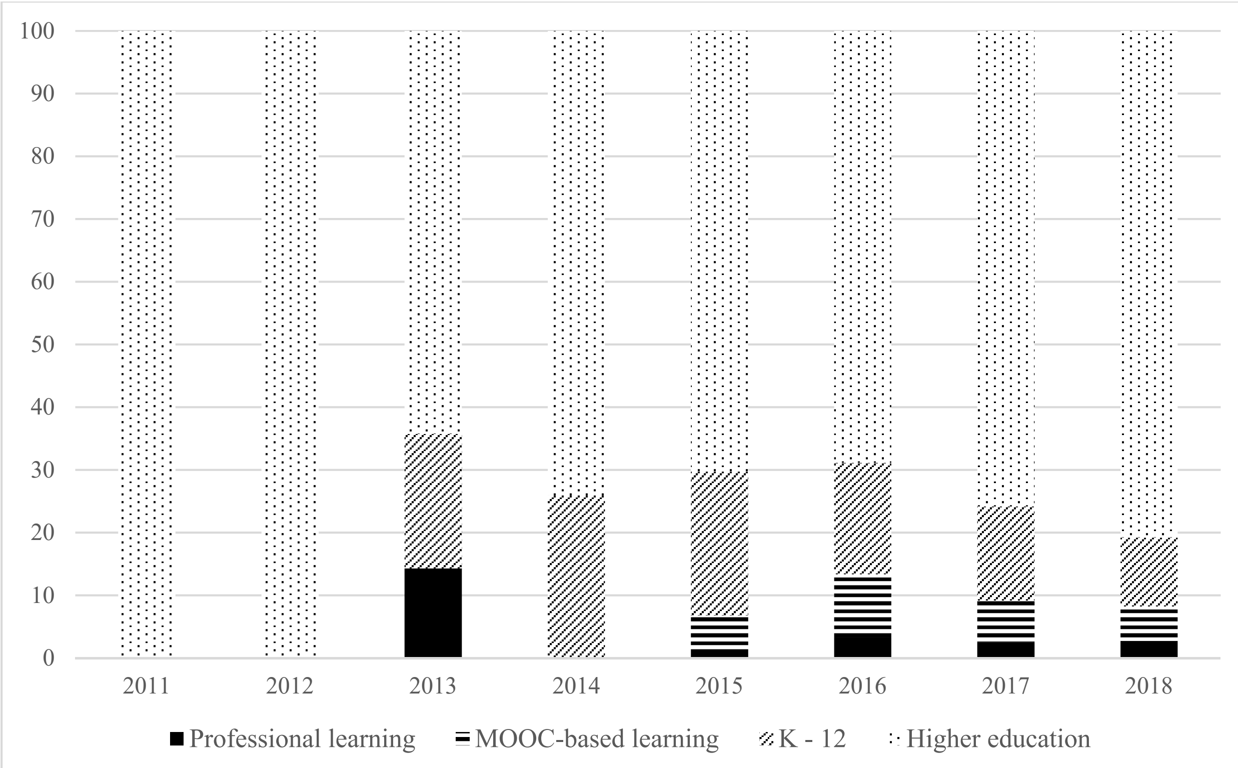

Figure 2. Changes in relative ratios of contexts in LA research reviewed

Research Question 3: What patterns exist in the methodologies used for research on LA?

We used a classification system of the SERVE Center (2008) to classify research methodologies that have been used for LA research. During the analysis of the research methodologies, we found articles that are hard to be classified by the classification system of the SERVE Center. Thus, we modified the classification system (see Table 5).

Table 5. Classification of research methodologies

Research methodology	Description
Descriptive-qualitative	Detailed descriptions of specific situation(s) using interviews, observations, and/or other qualitative data collection methods
Descriptive-quantitative	Numerical descriptions (frequency, average)
Correlation/regression/and other association analyses	Quantitative analyses of the strength of relationships between two or more variables using correlation analysis, regression analysis, likelihood ratio test, cluster analysis, random forest, path analysis, principle component analysis, analysis of variance, or other association analysis techniques
Quasi-experimental	Comparing an experimental group with a control group that is similar in characteristics but did not receive the intervention. Random assignment is not used to assign participants to an experimental group and a control group.
Experimental	Comparing an experimental group with a control group that is similar in characteristics but did not receive the intervention. Random assignment is used to assign participants to an experimental group and a control group.
Meta-analysis	Synthesis of results from multiple studies to determine the average impact of a similar intervention across the studies

Note. Adapted from the SERVE Center. (2008). Types of research methods. Retrieved from http://www.doe.virginia.gov/support/school_improvement/training/dta_student_support_sys/dropout_prevention/webinars_9-12/w2_s2_types_of_research_methods.pdf

Table 6 shows the frequency of research methodologies found in LA research reviewed. The results revealed that substantial research on LA employed the descriptive qualitative methodologies (n=254) or association analysis (n=200). Less research on LA employed experimental (n=47) or quasi-experimental (n=35) methodologies. *Figure 3* illustrates changes in the relative ratio of the methodologies in LA research reviewed.

Table 6. Frequency of research methodologies in LA research reviewed

Method	2011	2012	2013	2014	2015	2016	2017	2018*	Total
DQL	1	9	15	21	46	69	71	22	254
DQN	0	2	7	7	8	19	17	12	72
CRA	0	3	5	18	38	43	58	35	200
QEX	0	0	0	0	5	7	14	9	35
EXP	0	0	0	4	11	8	10	14	47
Total	1	14	27	50	108	146	170	92	608

Note. DQL = Descriptive-qualitative; DQN = Descriptive-quantitative; CRA = Correlational/regression analysis and other association analyses; QEX = Quasi-experimental; EXP = Experimental. The figures for 2018 are the number of research methodologies found in the articles published between January and July 2018.

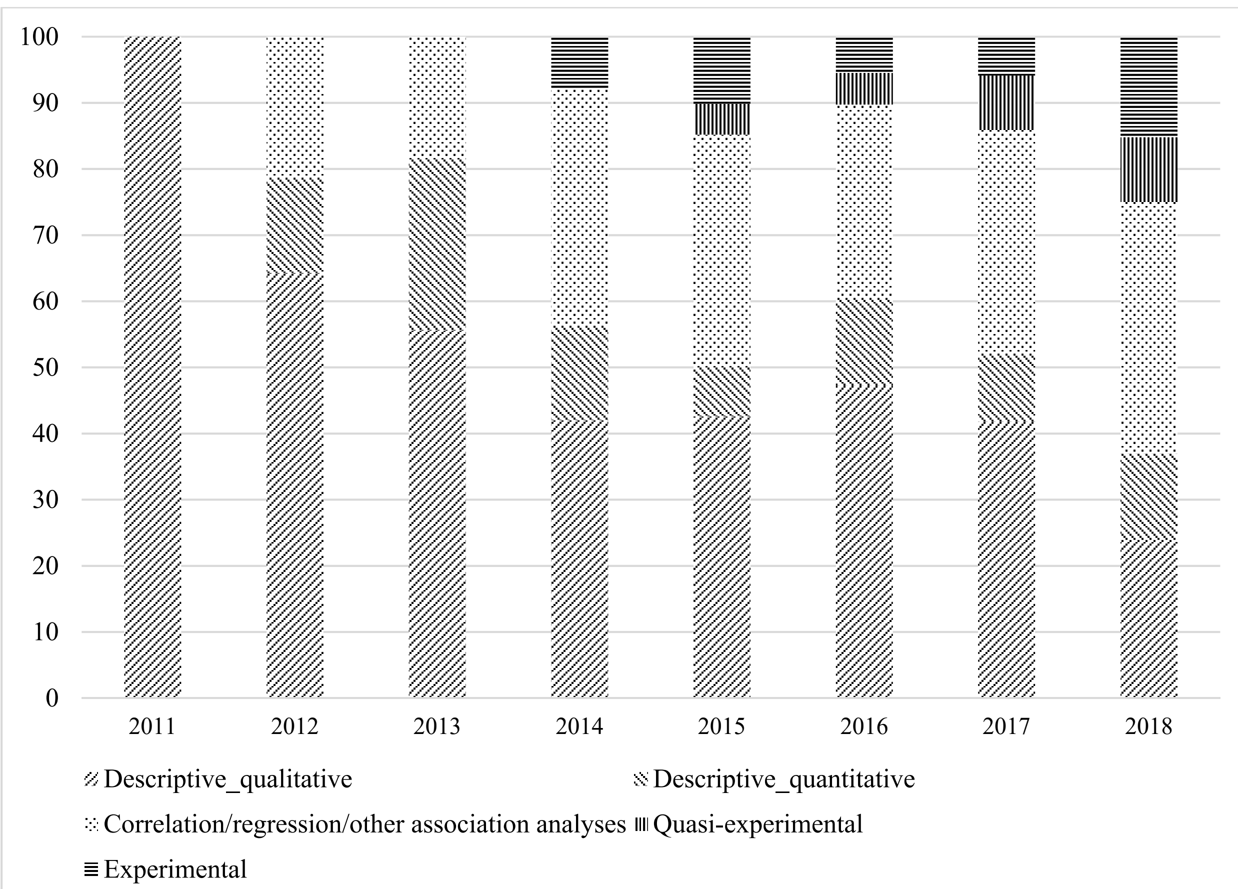

Figure 3. Changes in relative ratios of research methodologies in LA research reviewed

As shown in *Figure 3*, the majority of LA research (> 79% each year) between 2011 and 2013 employed descriptive methodologies, but the relative ratios of other research methodologies have steadily increased since 2016. In particular, research using experimental or quasi-experimental methodology has been increased noticeably.

Research Question 4: What patterns exist in the scopes of research on LA?

To classify the scopes of 608 research articles on LA, we developed an initial version of a classification framework based on our review of literature on the classification frameworks for LA or LA research. In developing the initial version of the classification framework, we included three layers: 1) Environment Layer, 2) Development Layer, and 3) Application Layer. During our review, however, we found scopes that are less relevant to these layers. Thus, we modified the classification framework and the final version of the classification framework consists of four layers: Foundation Layer, Environment Layer, Development Layer, and Application Layer. The Foundation Layer includes two scopes: 1) theoretical foundations of LA and 2) educational findings that can be used for LA. The Environment Layer includes three scopes: 1) legal and ethical environments, 2) technological environments, and 3) user's perception and behaviors. The Development Layer includes four scopes: 1) algorithm; 2) information presentation; 3) data collection, measure, and modeling; and 4) development methodology and process. Lastly, the Application Layer consists of four scopes: 1) descriptive LA, 2) diagnostic LA, 3) predictive LA, and 4) prescriptive LA.

Table 7 presents the frequency of the scopes in LA research reviewed. The results indicate that the three most studied scopes of LA research are user's perception and behaviors (n=102), theoretical foundations of LA (n=103), and educational findings that can be used for LA (n=108). On the other hand, the three least studied scopes of LA research are diagnostic LA (n=3), prescriptive LA (n=5), and predictive LA (n=12) in Application Layer.

Table 7. Frequency of scopes in LA research reviewed

Scope	2011	2012	2013	2014	2015	2016	2017	2018*	Total
APSC	0	0	0	2	0	0	2	1	5
APRD	0	0	0	2	3	1	3	3	12
ADGN	0	0	1	0	1	0	0	1	3
ADSC	0	1	0	2	11	6	13	7	40
DALG	0	0	2	2	18	15	18	21	76
DNPR	0	2	5	1	10	11	15	8	52
DDCM	0	2	3	3	9	9	13	5	44
DDMP	0	2	1	2	4	8	10	4	31
ELEE	0	0	1	2	1	15	2	2	23
ETEN	1	1	3	7	14	15	17	13	71
EUPB	0	2	2	8	16	25	29	20	102
FTFL	0	6	8	15	20	19	26	9	103
FEFL	0	1	3	7	14	38	33	12	108
Total	1	17	29	53	121	162	181	106	670

Note. APSC = Application Layer_Prescriptive LA; APRD = Application Layer_Predictive LA; ADGN = Application Layer_Diagnostic LA; ADSC = Application Layer_Descriptive LA; DALG = Development Layer_Algorithm; DNPR = Development Layer_Information Presentation; DDCM = Development Layer_Data Collection, Measure, Modeling; DDMP = Development Layer_Development Methodology and Process; ELEE = Environment Layer_Legal and Ethical Environment; ETEN = Environment Layer_Technological Environment; EUPB = Environment Layer_User's Perception and Behavior; FTFL = Foundation Layer_Theoretical Foundations of LA; FEFL = Foundation Layer_Educational Findings that can be used for LA. The figures for 2018 are the number of the scopes of research found in the articles published between January and July 2018.

Figure 4 illustrates changes in relative ratios of scopes in LA research reviewed. As shown in *Figure 4*, research dealing with Foundation Layer or Environment Layer has always counted for more than half of the scopes. Relatively less research studied the scopes in Development Layer, but the interest in these scopes has continued since 2012. Lastly, little research investigated Application Layer.

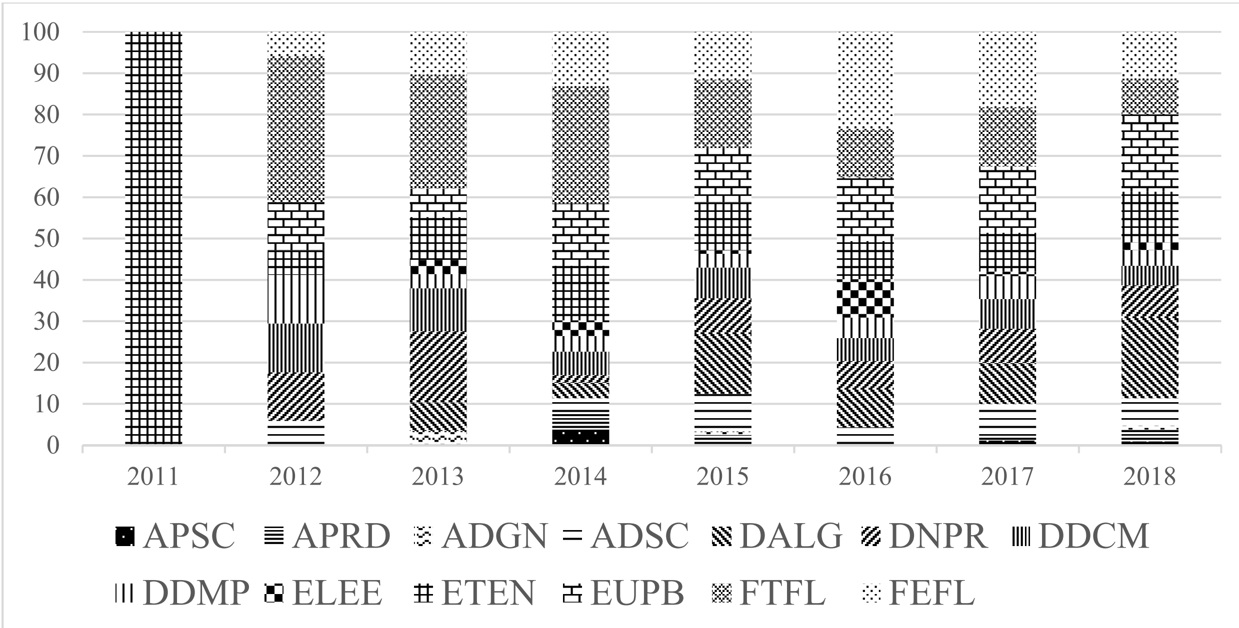

Figure 4. Changes in relative ratios of scopes in LA research reviewed

Research Question 5: How can the classification framework for LA research be developed?

After developing an initial version of a classification framework for LA research, we improved it to accommodate new types of research scopes in LA research. In improving the framework, we focused on flexibility and universality to adapt to and accommodate the theoretical and technological changes in future. *Figure 5* illustrates the classification framework for LA research. To validate the classification framework, we classified 670 scopes found from 608 articles into four layers and 13 types.

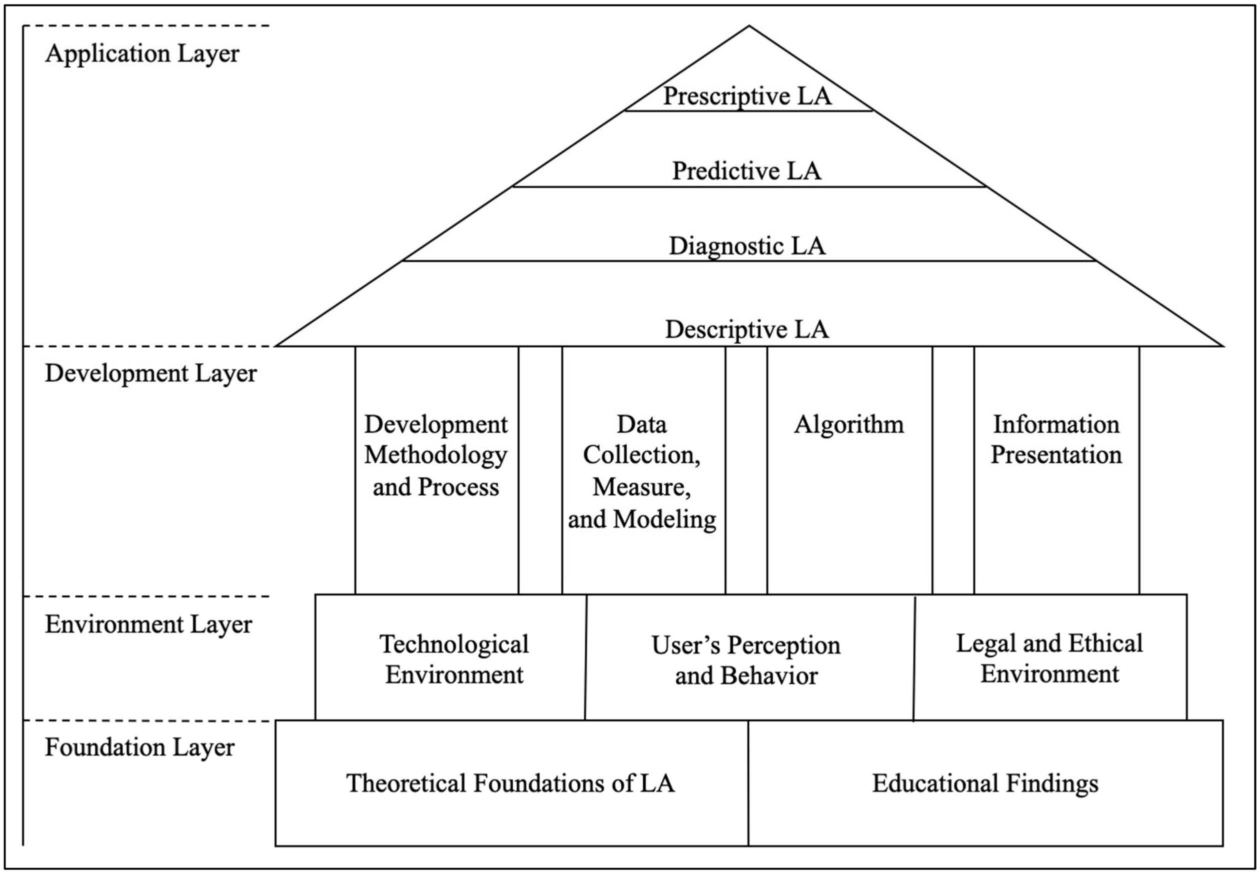

Figure 5. Classification framework for LA research

Research Question 6: What are some gaps in research on LA for professional learning?

We mapped each of the research articles on LA for professional learning (n=11) with the classification framework for LA research. The results revealed that there are many research gaps in LA for professional learning (see Table 8). No research included in our review studied predictive LA, diagnostic LA, descriptive LA, algorithms for LA, development methodology and process, legal and ethical environments, and theoretical foundations of LA for professional learning.

Table 8. Frequency of scopes in research on LA for professional learning

Classification of LA research	Frequency of scopes
Application Layer_Prescriptive LA	1
Application Layer_Predictive LA	0
Application Layer_Diagnostic LA	0
Application Layer_Descriptive LA	0
Development Layer_Algorithm	0
Development Layer_Information Presentation	1
Development Layer_Data Collection, Measure, Modeling	2
Development Layer_Development Methodology and Process	0
Environment Layer_Legal and Ethical Environment	0
Environment Layer_Technological Environment	1
Environment Layer_User's Perception and Behavior	3
Foundation Layer_Theoretical Foundations of LA	0
Foundation Layer_Educational Findings that can be used for LA	3

Discussions

Definition of LA

We found 31 different definitions of LA from our review. Though there are some differences in the definitions of LA, the definitions share a few commonalities. However, there is a big difference between disciplines in defining LA. While researchers in the field of education take a broad approach to defining LA, ones in the field of engineering and business take a narrow approach. The former include analyzing data by human or a tool separated from a system that stores target data in LA research, but the latter consider LA research as one studying an automated tool that collects, analyzes, and provides data or information. Ifenthaler and Widanapathirana (2014) used the term 'engine' to mean the functions of the automated tool, but it is not sufficient to reflect the different views from the various disciplines. It seems there is a need to use the broad definition and narrow definition separately.

Another issue regarding the definition of LA is that the most frequently used definition, the definition of Society for Learning Analytics Research (2011), has the following problems:1) It includes 'measurement' though measurement is not necessary for LA (Knight & Littleton, 2015) and 2) it includes 'optimizing' though the optimizing can be interpreted in different ways. At this point, thus, it seems there is a need to redefine LA to clarify its scope and purpose.

We propose a broad definition and a narrow definition of LA. For example, a broad definition of LA refers to collection, analysis, and reporting of data or information on learners and their learning experiences to understand and improve learning. A narrow definition of LA refers to an automated system that collects, analyzes, and reports data or information on learners and their learning experiences to provide information supporting decision making regarding learning.

Research context

There is a significant imbalance in LA research in favor of the context of higher education. We assume that the imbalance probably occurred due to the relative ease of accessing data within the context. Thus, it seems necessary for researchers to collaborate with the stakeholders who are in the under-researched contexts. For example, researchers in the field of LA can study LA for professional learning more actively by collaborating with stakeholders in corporations.

Research methodology

Though there seemed to have been an imbalance of LA research in terms of research methodology in the early years of LA, the imbalance seems to be resolved by the researchers who employed experimental or quasi-experimental methodologies for their LA research. Comparing the contexts and scopes of LA research, the research methodology in LA research seems balanced.

Research scope

There is a significant imbalance in LA research in terms of the research scope. While considerable research dealt with the scopes in Foundation Layer and Environment Layer, little research covered the scopes in Application Layer. Thus, researchers in the field of LA can find relatively many opportunities for research.

In addition, in the field of Multimodal Learning Analytics (MMLA), researchers have been interested in using a variety of types of data that can be used as input sources for LA. Some researchers (Abrahamson, Shayan, Bakker, & van der Schaaf, 2015; Lau et al., 2018; Lu, Zhang, Zhang, Xiao, & Yu, 2017; Munoz et al., 2018; Prieto, Sharma, Kidzinski, Rodríguez-Triana, & Dillenbourg, 2018; Zaletelj & Košir, 2017) used data from sensors; such as electroencephalogram (EEG) sensors, motion sensors, and eye-tracking sensors; as data to be analyzed by LA engines. However, the multimodality can be applied for reporting information as well as collecting data. For example, future research on MMLA may focus on application of text-to-speech technologies and sonification to report information.

Lastly, researchers might benefit from expanding the types of collected data for LA. Currently, many researchers in the field focus on LMS data on learners' activities. However, this approach will limit the understanding of learners' learning experiences as there can be many other factors, such as learning materials, content of instructor's feedback, and contents of conversations between peer learners, influencing the learners and their learning experiences. Thus, expanding the types of collected data for LA seems a good approach. Such expanded data types include social learning analytics (de Laat & Prinsen, 2014; Shum & Ferguson, 2012), learners' gestures (Viswanathan & VanLehn, 2018; Zaletelj & Kosir, 2017), eye movement tracking (Abrahamson, Shayan, Bakker, & van der Schaaf, 2015), and electroencephalography (Lau et al., 2018).

Classification framework for LA research

We developed the classification framework with a focus on flexibility and universality to respond to possible changes in regard to theory and technology in the future. However, other researchers in the field may find some chances to improve the framework. For example, the scopes in the Foundation Layer of the framework could be divided into more detailed scopes or add newly discovered scopes.

For this review, we included articles only from academic databases. It might be meaningful to use Google Scholar, a search engine for scholarly literature, to include more research articles to validate our classification framework for LA research.

LA research for professional learning

As shown in Table 8, there are many research gaps in LA for professional learning. There might be many reasons regarding such gaps, but we assume that the relative ease of accessing higher education data for research is one of the most influential factors on this phenomenon. To resolve this, as mentioned earlier, it is necessary for researchers to collaborate with the stakeholders who are responsible for professional learning in various organizations.

In addition, to promote research on LA for professional learning, studying relationships and interactions between LA and other systems, such as talent management system or succession planning system. Based on the understanding of the relationships and interactions, researchers may find more research opportunities.

Limitations

Though we conducted iterative analyses for classifying articles that couldn't be clearly classified into a category, there can be a different view for the classification. Therefore, a further discussion needs to be made with regard to the validation of the classification framework for LA research we developed.

References

Abrahamson, D., Shayan, S., Bakker, A., & van der Schaaf, M. (2015). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development, 58*(4-5), 218-244. doi:10.1159/000443153

Aljohani, N. R., Daud, A., Abbasi, R. A., Alowibdi, J. S., Basheri, M., & Aslam, M. A. (2019). An integrated framework for course adapted student learning analytics dashboard. *Computers in Human Behavior*, *92*, 679-690.

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*(4), 564-599. doi:10.1080/10508406.2013.836655

Brown, M. (2011). Learning analytics: The coming third wave. *EDUCAUSE Learning Initiative Brief*. Retrieved from https://library.educause.edu/-/media/files/library/2011/4/elib1101-pdf.pdf

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE review, 42*(4), 40.

Cooper, A. (2012). What is analytics? Definition and essential characteristics. CetisAnalytics Series, 1(5). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.269.7294&rep=rep1&type=pdf

de Laat, M., & Prinsen, F. R. (2014). Social learning analytics: Navigating the changing settings of higher education. *Research & Practice in Assessment, 9*, 51-60.

Fleckenstein, M., & Fellows, L. (2018). Modern data strategy. Cham, Switzerland: Springer. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=1709311&site=eds-live&scope=site

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4(5/6), 304-317.

Goldstein, P. J. (2005). Academic analytics: The uses of management information and technology in higher education. ECAR Key Findings. Retrieved from https://er.educause.edu/-/media/files/articles/2007/7/ekf0508.pdf

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. Health Information & Libraries Journal, 26(2), 91-108.

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

Hicks, C. (2018). Predicting knowledge workers' participation in voluntary learning with employee characteristics and online learning tools. Journal of Workplace Learning, 30(2), 78-88. doi:10.1108/jwl-04-2017-0030

Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. Technology, Knowledge and Learning, 19(1-2), 221-240.
Knight, S., & Littleton, K. (2015). Dialogue as data in learning analytics for productive educational dialogue. *Journal of Learning Analytics, 2*(3), 111-143.
Kumar, R. (2017). Machine learning and cognition in enterprises: Business intelligence transformed. Berkeley, CA: Apress. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=1636824&site=eds-live&scope=site
Lau, K. H. V., Farooque, P., Leydon, G., Schwartz, M. L., Sadler, R. M., & Moeller, J. J. (2018). Using learning analytics to evaluate a video-based lecture series. *Medical Teacher, 40*(1), 91-98. doi:10.1080/0142159x.2017.1395001
Lawson, C., Beer, C., Rossi, D., Moore, T., & Fleming, J. (2016). Identification of 'at risk' students using learning analytics: The ethical dilemmas of intervention strategies in a higher education institution. *Educational Technology Research and Development, 64*(5), 957-968. doi:10.1007/s11423-016-9459-0
Long, P., Siemens, G., Conole, G., & Gašević, D. (2011). Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK11), Banff, AB, Canada, Feb 27-Mar 01, 2011. New York: ACM.
Lu, O. H. T., Huang, A. Y. Q., Huang, J. C. H., Lin, A. J. Q., Hiroaki, O., & Yang, S. J. H. (2018). Applying learning analytics for the early prediction of students' academic performance in blended learning. Journal of Educational Technology & Society, 21(2), 220-232.
Lu, Y., Zhang, S., Zhang, Z. Q., Xiao, W. D., & Yu, S. Q. (2017). A framework for learning analytics using commodity wearable devices. *Sensors, 17*(6), 25. doi:10.3390/s17061382
Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. Computers & Education, 103, 1-15. doi:10.1016/j.compedu.2016.09.005
Martín-Monje, E., Castrillo, M. D., & Mañana-Rodríguez, J. (2018). Understanding online interaction in language MOOCs through learning analytics. Computer Assisted Language Learning, 31(3), 251-272. doi:10.1080/09588221.2017.1378237
Munoz, R., Villarroel, R., Barcelos, T. S., Souza, A., Merino, E., Guinez, R., & Silva, L. A. (2018). Development of a software that supports multimodal learning analytics: A case study on oral presentations. *Journal of Universal Computer Science, 24*(2), 149-170.
Muslim, A., Chatti, M. A., Bashir, M. B., Varela, O. E. B., & Schroeder, U. (2018). A modular and extensible framework for open learning analytics. Journal of Learning Analytics, 5(1), 92-100.
Peña-Ayala, A. (2018). Learning analytics: A glance of evolution, status, and trends according to a proposed taxonomy. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(3), e1243.
Prieto, L. P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M. J., & Dillenbourg, P. (2018). Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning, 34*(2), 193-203. doi:10.1111/jcal.12232
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
SERVE Center. (2008). Types of research methods. Retrieved from http://www.doe.virginia.gov/support/school_improvement/training/dta_student_support_sys/dropout_prevention/webinars_9-12/w2_s2_types_of_research_methods.pdf
Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Journal of Educational Technology & Society, 15*(3), 3-26.
Skourletopoulos, G., Mastorakis, G., Mavromoustakis, C. X., Dobre, C., & Pallis, E. (2018). Mobile big data: A roadmap from models to technologies. Cham, Switzerland: Springer. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=1632634&site=eds-live&scope=site
Society for Learning Analytics Research. (2011). *Call for papers of the 1st International Conference on Learning Analytics & Knowledge (LAK 2011)*. Retrieved from https://tekri.athabascau.ca/analytics/call-papers
Siemens, G. (2010). *What are Learning Analytics?* Retrieved from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/
Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*, 1510-1529
van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. EDUCAUSE learning initiative, 1(1), l-ll.
Viswanathan, S. A., & VanLehn, K. (2018). Using the tablet gestures and speech of pairs of students to classify their collaboration. *IEEE Transactions on Learning Technologies, 11*(2), 230-242. doi:10.1109/tlt.2017.2704099

Willis, J., Zilvinskis, J., & Borden, V. M. H. (2017). An overview of learning analytics. *New Directions for Higher Education, 2017*(179), 9-17.

Yassine, S., Kadry, S., & Sicilia, M. A. (2016, April). A framework for learning analytics in moodle for assessing course outcomes. In 2016 IEEE Global Engineering Education Conference (EDUCON) (pp. 261-266). IEEE.

Yu, L. C., Lee, C. W., Pan, H. I., Chou, C. Y., Chao, P. Y., Chen, Z. H., . . . Lai, K. R. (2018). Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. Journal of Computer Assisted Learning, 34(4), 358-365. doi:10.1111/jcal.12247

Zaletelj, J., & Kosir, A. (2017). Predicting students' attention in the classroom from Kinect facial and body features. *EURASIP Journal on Image and Video Processing, 2017*:80. doi:10.1186/s13640-017-0228-8