

Learning Analytics for Blended Learning

A Systematic Review of Theory, Methodology, and Ethical Considerations

<https://doi.org/10.3991/ijai.v2i2.17887>

Nina Bergdahl ^(✉), Jalal Nouri, Thashmee Karunaratne, Muhammad Afzaal
Stockholm University, Stockholm, Sweden
ninabe@dsv.su.se

Mohammad Saqr
KTH Royal Institute of Technology, Stockholm, Sweden

Abstract—Learning Analytics (LA) approaches in Blended Learning (BL) research is becoming an established field. In the light of previous critique toward LA for not being grounded in theory, the General Data Protection and a renewed focus on individuals' integrity, this review aims to explore the use of theories, the methodological and analytic approaches in educational settings, along with surveying ethical and legal considerations. The review also maps and explores the outcomes and discusses the pitfalls and potentials currently seen in the field. Journal articles and conference papers were identified through systematic search across relevant databases. 70 papers met the inclusion criteria: they applied LA within a BL setting, were peer-reviewed, full-papers, and were written in English. The results reveal that the use of theoretical and methodological approaches were disperse. We identified approaches of BL not included in categories of BL in existing BL literature and suggest these may be referred to as complex hybrid learning, and that ethical considerations and legal requirements often have been overlooked. We highlight critical issues that contribute to raise awareness and inform alignment for future research to ameliorate diffuse applications within the field of LA.

Keywords—Literature review; learning analytics, blended learning, complex hybrid learning

1 Introduction

1.1 The emergence of learning analytics

Given the wealth and complexity of learning, learning sciences have become an interdisciplinary domain that includes cognitive, educational, social and computer science among others. Ten years ago, learning analytics (LA) emerged as a multidisciplinary research domain with the overarching premise to harness the power of data and analytics to advance our understanding of learning as well as help improve learning,

teaching and optimize the learning environments [26]. As such, 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” [83]. Interest in LA was catalysed by three main factors; firstly: the rewarding business intelligence and big-data success stories that has contributed to industry growth and added competitive advantage to companies by allowing them to better understand customers and offer better recommendations (ibid.). Secondly, the availability of immense volumes of digital traces and clickstream data recorded by learning management systems and other digital learning environments, such as student’s information systems, online library platforms and video streaming services. Third, the revolutionary developments in data science methods and improved computer hardware that became more powerful and accessible [15, 21, 83]. Inspired by the industry, LA has initially used digital trace data to create predictive models to for example, forecast drop-outs, identify students at-risk of failing, or offer visual dashboards. However, criticism had been levied at these models for failing to account for contextual factors, being atheoretical and difficult to replicate [25, 100]. Research has since extended both data collection, analysis methods and approaches to theory in LA research. Recently, data collection methods have grown in volume and diversity to cover the full breadth of learners’ activities e.g., classroom interactions, physiological indices, proximity data, eye tracking, self-reports in addition to the commonly used digital traces. Similarly, methods have grown to include sequence and process mining, epistemic network analysis as well as advanced machine learning algorithms [16, 20, 62].

Today, when it is common that students’ educational reality is an integration of physical and virtual learning environments, often referred to as a Blended Learning (BL) [37-39], the data collection methods are often broadened to survey both the physical and the virtual spaces. Applying such data collection and theoretically aligned refined models, researchers hope to capture learners’ behaviour where it occurs. The study of LA and BL is a growing field of inquiry [48], which is garnering a growing attention in and outside the LA community. The rationale for undertaking this review is that LA research has been criticised for not being grounded sufficiently in theory. For example, that perspectives of learning are lacking [57], and that BL research have remained vague and unclear, why there has been a call for research to further develop definitions of BL, models and conceptualisations [46]. Moreover, as LA is becoming an established research field of its own, it is interesting to explore the methodological practices applied across LA research and with the fast-paced development of big data analytics on individual trace data and the implementation of the General Data Protection Regulations of the European Union (GDPR), valid concerns might be raised with regards to if and how ethical and legal considerations have been applied. Accordingly, a systematic review that identify theoretical underpinnings, methodological practices, considerations of ethical aspects and legal requirements and the contributions in LA research is needed to raise awareness and inform alignment for future research.

2 Background

2.1 Learning analytics in blended learning

Blended Learning is a term coined in the 90's, which related practices over the years gained substantial influence, and today is regarded as the “new normal” [46]. The concept of BL, its operationalisations and definitions are wide and still evolving [1]. Researchers have highlighted that BL is a broad term that may reflect different variations to what is blended, the extent and duration of the blend, models of blended learning, a systematic approach that include any part of a system that combines face-to-face (f2f) instruction with instructions mediated by digital technologies [37, 46] which commonly include blended instruction, blended delivery, blended learning environments [37] spatiotemporal aspects, where learning can be self-paced and individual qualitative aspects [56] which reflect thoughtful integration [31]. In short, BL can be viewed as an umbrella term, that without specific descriptions will not inform the reader what aspects of teaching and learning that is approached.

In addition, LA approaches has been critiqued for being atheoretical [32, 57]. In LA and BL research, as in other educational research, the main aim is to support students to succeed with their education. Such effort can be demonstrated by including theory that provides guidance for how to understand, operationalise, measure and interpret for example students' engagement in learning. Engagement theories may emphasize different aspects, for example agency [5] or cognition [6] related to self-regulation (SRL). Student engagement is critical for learning, and from this perspective LA in BL is warranted as it combines the BL setting with theoretical insights of students use of digital technologies, their ability to self-regulate to re-engage in the face of difficulty, distraction, frustration, simultaneous social demands et cetera. However, critique has forwarded that LMS data may not be suitable to capture a nuanced understanding of student engagement, this as engagement is a multi-dimensional construct, and LMS data, still, at best can reflect a one-dimensional aspect of engagement, [42]. Moreover, researchers [42] could not find any significant correlation between student self-reports of engagement and the LMS trace data. Thus, if the approaches and applications of BL, theories from learning perspectives and how these are operationalised are lacking, and self-reports and traces data are not correlating, this decreases the value of the contributions in the field of LA. Previous reviews have surveyed theoretical underpinnings in LA research, and concluded that the grounding in (educational) theory is evident but too often meagre or lacking [e.g. 57, 96]. For example [57] concluded that existing learning analytics dashboards research are rarely grounded in learning theory, cannot be suggested to support metacognition and thus, do not inform teachers of effective practices and strategies. LA reviews have also explored methods applied within LA research [e.g. 6] and identified that LA studies use diverse methods, for example, visual data analysis and clustering techniques, social network analysis (SNA) and educational data mining. Taken together, however, the existing reviews on LA research have not taken contributions into account, such approach is critical, as if applications of BL, theories from learning perspectives and how these are operationalised are insufficient or lacking, the contributions becomes unclear.

Today, there are a substantial number of Learning Analytics reviews. LA reviews often specialise on particular areas like for instance: game learning analytics [9], visual learning analytics, [96], the role of self-regulated learning (SRL) [98], learning analytics dashboards [57], or the uses of LA in relation to specific methods or approaches e.g.; open learner models [10] educational data mining [12] or apply a wider scope that explore national research efforts, policies, infrastructures and competence centres across several European countries [68].

While several reviews highlight similar findings (i.e. a lack of theoretical underpinning, unclear uses of methods) there is a risk of transferring and projecting findings across LA research, as findings which might not reflect the broader LA research, which in turn may lead to overgeneralisations. Although there are many published (systematic, scoping and area-specific) reviews on LA in online settings, in order to understand their aim, objective and contribution, it is beneficial to approach a less specific overview of LA research to survey commonalities, of theoretical underpinnings (including conceptualisations of BL and learning perspectives) methodological approaches, ethical and legal requirements and contributions.

However, in addition to theoretical and methodological aspects, an additional layer of complexity is added to LA research in a BL environment. LA is in itself a practice of gathering, analysing and sharing big amounts of personal data, which comes with an increased need for ethical considerations and adherences to legal requirements. The ethical, privacy and legal concerns of processing of personal data are on the frontier of data processing due to the presence of the GDPR [14]. LA is a subject developed on data-driven approaches to education innovation, and hence, in the spotlight of this concerns. Beyond ethics, the GDPR provides a legal framework in preserving the rights of the data subjects, that is: the students. Learning analytics operates on data about students and their learning environments, where personal data of the students is an integral part. Personal data of students refers to any data that directly or indirectly connected to an identifiable person, e.g., student names, personal identification numbers, email, photographs, and other data that could lead to identifying an individual [29]. It is typical that learning and student management systems store, retrieve, and process such data, driven by different academic and learning purposes [15]. While the absence of ethical considerations [16], [17], privacy issues and GDPR [18] have been previously critiqued in regards to the adoption of LA we did not find any existing review that had explored these aspects of GDPR, ethics and privacy on LA research. Therefore, in this study, we have added a focus on how the reviewed studies consider ethical and legal aspects of using data. Informed by these previous concerns and critiques we raised the following questions:

1. How is blended learning defined in the reviewed learning analytics research?
2. For which learning focus perspectives are theories used in the reviewed learning analytics research?
3. What approaches of data collection, methods and analysis are evident in the reviewed learning analytics research?
4. How are ethical and legal aspects considered in the reviewed learning analytics research?
5. What are the contributions of the reviewed learning analytics research?

3 Method

3.1 Search strategy and selection procedure

This study examines academic journals and conference papers applying Learning Analytics in Blended Learning from two databases (see Table 1). A systematic search was conducted using EBSCOhost via Stockholm University library (filtered by content providers: Scopus, ERIC, Academic Search Premier, Directory of Open Access Journals, ScienceDirect) for academic journals, and ACM DL Digital Library, for conference papers. As detailed below, the systematic search via EBSCOhost followed educational journals by status [13], and the selection employed journal rankings provided by SCIMAGO Institutions Rankings.

Table 1. Overview of search string results

Database and journal search strings	Hits
EBSCOhost database search string	
“learning analytics” + “blended learning”	79
“learning analytics” + “blended learning” (incl. “within full text of articles”)	282
“learning analytics” + “blended environment”	0
“learning analytics” + “blended learning environment”	2
“teaching analytics” + “blended /”-learning”/ “-environment”	0
“educational data mining” + “blended learning”	8
“educational data mining” + “blended”	8
“educational data mining” + “blended environment”	0
ACM DL Digital Library	
“learning analytics” + “blended learning” + “LAK” (Learning Analytics & Knowledge conference)	21
“educational data mining” + “blended learning” + “LAK” (Learning Analytics & Knowledge conference)	11
“learning analytics” + “blended learning”	43
“educational data mining” + “blended learning”	22
Journal search via EBSCOhost via Stockholm University Library and the Journal of Learning Analytics	
“blended learning”	7
“learning analytics” + “blended learning”	6
“educational data mining” + “blended learning”	3
Internet and Higher Education	
“learning analytics” + “blended learning”	14
“educational data mining” + “blended learning”	8
Educational Technology and Society	
“learning analytics” + “blended learning”	2
“educational data mining” + “blended learning”	0
Journal of Computer Assisted Learning	
“learning analytics” + “blended learning”	2
“educational data mining” + “blended learning”	0
British Journal of Educational Technology	
“learning analytics” + “blended learning”	16

“educational data mining” + “blended learning”	5
Computers in Human Behavior	
“learning analytics” + “blended learning”	26
“educational data mining” + “blended learning”	17
Computers and Education, Communications in Information Literacy, Learning and Instruction, International Review of Research in Open and Distance Learning, Educational Evaluation and Policy Analysis International Journal of Mobile and Blended Learning	
“learning analytics” + “blended learning”	0
“educational data mining” + “blended learning”	0

The search combinations used in SCIMAGO: Social Sciences + E-learning + All regions / countries + Journals + 2017; Social Sciences + Education + All regions / countries + Journals + 2017; Computer Science + Human-Computer Interaction + All regions / countries + Journals + 2017; and Computer Science + Human-Computer Interaction + All regions / countries + Journals + 2017. Inclusion from each 4 search combinations above was determined by relevance of the title and the choice was limited to the top-10 journals in each search combination. We identified papers from the following six journals Internet and Higher Education, Journal of Computer Assisted Learning, British Journal of Educational Technology, Computers in Human Behaviour, Educational Technology and Society, and the Journal of Learning Analytics. To search combinations used for the EBSCOhost database, we used combinations of keywords: “learning analytics” + “blended learning”; “learning analytics” + “blended environment”; “teaching analytics” + “blended learning”; “teaching analytics” + “blended environment”; “educational data mining” + “blended learning”; “educational data mining” and “blended environment”. We included peer-reviewed, academic journals, written in English. We also tried including a “search within full text of articles”; and screened titles and abstracts of the papers for inclusion, and remove duplicates. We decided to not utilise the function further, as it returned irrelevant articles where BL and LA were mentioned only in the reference section. We searched for articles published the between January 2013-July 2020.

Overall, the keyword searches amounted to 304 hits (not including the search within full text of articles). After removing the duplicates, 193 journal articles and conference papers remained; 38 hits did not return full texts and 4 hits returned hits in other languages (three in Danish and one in German) although the search criteria were aiming at English texts only. After that, we sifted through the remaining papers and excluded 32 papers that were not directly relevant to LA and BL, and 49 that lacked one of the two focuses (either LA or BL). During close-reading, an additional three papers were excluded, as they did not meet the inclusion criteria. Thus, which we proceeded to code and later analyse 70 papers.

3.2 Data coding and analysis

Following a coding scheme all articles were read through by two authors, who sorted the content in: article data (country, publication year, title), educational context, (blended learning interpretation and level), research aims/questions, theoretical

underpinnings and definitions of BL, data sources, data collection methods, ethical considerations, analytical methods and results and contributions. All authors then conducted a deeper analysis of one section of the reviewed articles each (1. theoretical underpinnings, 2. data collection, methods and analysis, 3. ethical and legal considerations and 4. contributions). In depth discussions were held between the authors to discuss approaches and align findings.

4 Results

The result section details the findings as follows: 4.1 Theoretical underpinnings, 4.2 Data collection, methods and analysis, 4.3 Ethical and legal considerations and 4.4 Contributions of the reviewed articles.

4.1 Theoretical underpinnings

To discern the positioning of the articles in terms of their relation to BL, we analysed the articles with regards to how blended learning was used throughout the articles, in particular how frequent the authors refer to blended learning, their definition, description and use of theory. Currently, BL literature [e.g. 37, 39] have identified three common ways in which explorations of blended learning delivery may vary: blended instruction, blended distribution, as identified in [14, 15] or blended pedagogies [54]. However, going through these descriptions, we also found that studies could displayed a combination of blended instruction and blended distribution; i.e. when a section of the course is provided fully f2f, followed by the remainder offered fully online [79] or reversed: a course is delivered fully online and then fully f2f [49]. We also identified that the BL was used in ways beyond these categories. We identified a combination of blended learning approaches in which some, but not all, students use the BL component. For example, we identified studies that offer i) optional adoption of the blended component to the students, [e.g. 43, 80, 87], or when ii) there was a synchronous teaching of f2f students and online students in the same classroom [9] and iii) in cases of reversed distribution; a channel directed exclusively from the student to peers and/or teachers, in which the teaching (distribution, delivery and pedagogy) has remained traditional, for example as an e-portfolio accessible in a social network [35] or in flipped classrooms, where students responded to distributed (asynchronous) media and instruction in their own time and place. [40].

Table 2. Blended learning definitions

Blended learning definitions	
“Technology to support face-to-face teaching and to enhance student participation” (Liao & Lu, 2008).	[2]
“Blended learning system as one which combines face-to-face instruction with computer-mediated instruction with the aim of complementing each other” (Graham, 2006; 2009; 2013)	[7] [32] [58][63] [75][102]
“The range of possibilities presented by combining Internet and digital media with established classroom forms that require the physical co-presence of teacher and students” (Friesen, 2012)	[23]
“B-learning is the form of learning environment where the traditional classroom teaching and face-to-face communication between teacher and students are blended with the computer-mediated interaction “(Bubaš & Kermek, 2004)	[30]
“Blended learning is a combination of traditional face to face learning and online learning. It has the advantages of the both, providing students with unique flexible learning experience and becoming one of the fastest growing trends in educational field” (Thorne, 2003)	[36]
“The thoughtful integration of classroom face-to-face learning experiences with online learning experiences” (Garrison & Kanuka, 2004)	[41] [76]
“Taking the best from self-paced, instructor-led, distance and classroom delivery to achieve flexible, cost-effective training that can reach the widest audience geographically and in terms of learning styles and levels” (Marsh & Drexler, 2001)	[44]
“The integration of thoughtfully selected and complementary face-to-face and online approaches and technologies” (Garrison & Vaughan, 2008)	[60]
“Blended learning is learning that happens in an instructional context which is characterized by a deliberate combination of online and classroom- based interventions to instigate and support learning. Learning happening in purely online or purely classroom-based instructional settings is excluded” (Boelens, Van Laer, De Wever & Elen, 2015).	[94]

Table 2 shows an overview of the used definitions of blended learning. While 29% of the articles offered a clear definition, most articles relied on inferences or contextual descriptions. 18% of the articles neither inferred nor described BL. The articles that offered a definition most commonly cited Graham [37-39]

Analysis from a learning focus perspective: Revealed five themes reflecting the perspective of the research: (i) the flipped classroom, (ii) collaborative learning, (iii) conversational aspects of learning, (iv) engagement and self-regulation operationalised using system trace data and (v) learner profiles and procrastination. Studies that include theories are presented in a condensed and summarised form (the others are not).

1. **The flipped classroom:** While most studies exploring the flipped classroom, approached student engagement and learning, a few were focusing on the actual learning situation [19–21]. These studies applied a more over-arching, abstract level of theory to inform their study, and also discussed their findings in the light of theory. However, while SRL were, by far, the most commonly used theory to explore flipped classroom design, most studies did not seek to explore the blended learning environment.
2. **Collaborative learning:** Social Network Analysis was used to visualise online interactions, and identify productive behaviours and correlation with performance [35, 41, 43, 81]. These used constructivist and situated learning theories and theories of self-regulation.

3. **Conversational aspects of learning:** Studies exploring conversational aspects of learning, most commonly approached feedback operationalised as online reports, referring to feedback and assessment theories [e.g. 72, 90], or deep learning theories [40, 51]. Another type of input to learning was explored by [89], grounded their study in the Dispositional Learning Analytics (DLA) infrastructure, used previous publications on assistant conversational agents, and theories on cognitive load in microblogging. Using the foundation of the Community of Inquiry framework, which prioritises teacher presence, and active participation, [79] used trace data to operationalise active participation, as the number of: messages sent, documents uploaded, chat sessions attended, as well as data collected to analyse teacher presence.
4. **Engagement and self-regulation operationalised using system trace data:** Out of all the theories applied, engagement in general and self-regulated learning (SRL) in particular, were the most commonly used. To add to these research approaches, aspects of culture and gender were introduced and explored [86]. While SRL often was operationalised as observable indicators in system logs, motivation was approached by measures of self-efficacy, intrinsic value, test anxiety, cognitive strategy and self-regulation using a questionnaire. SRL was often operationalised as trace data, and combined with other engagement and learning theories. [36, 52, 58, 80]. Numerous studies explored relations between trace data, performance and SRL using self-reports [30]; some in combination with other theories, for example theories of motivation [20] socio-cultural perspectives [31], Self-Determination Theory and the Control-Value Theory of achievement emotions [86-88].
5. **Learner profiles:** We identified that with studies exploring learner profiles, it was common to inform this approach with other theories, for example, course satisfaction and social constructivist theory [18], deep and shallow learning [32] active learning and engagement [35] and procrastination [1, 54, 66]. In the reviewed studies, student learning strategies was often operationalised as trace data on student interaction with online learning resources [33]. Amongst these, procrastination was found to be common. Several studies operationalised SRL as procrastination [54, 66]. For example, using LMS data to survey time spent studying and time spent refraining from accessing available data [54]. Procrastination was also explored without relation to SRL, or how long the student waited before accessing LMS materials [1]. Other researchers used questionnaires to survey procrastination and risk taking using the Expectancy-Value Theory, motivation, using the Academic Motivational Scale and help seeking, and epistemic emotions to inform a to approach how different learning strategies relate to preferences of feedback [66].

In sum: While most reviewed studies approaching a flipped classroom, used theories with a focus on students and their engagement and learning, a few were focusing on the actual learning situation [69, 75, 76, 94] or combined flipped classroom theories of Computer Assisted Language-Learning [34]. The latter studies applied a more overarching, abstract level of theory to inform their study, and also discussed their findings in the light of theory. Some studies argued that there is a need to develop a specific SRL-LA theory [63]. However, while SRL were, by far, the most commonly used

theory, most studies did not seek to explore the blended learning environment, but seemed to relate their data collection to operationalisations related to a learning perspective, with or without underpinning theories of learning.

4.2 Data collection, methods and analysis

All studies included in the review used a digital platform for collecting data. As can be expected, LMS was the most used platform for data collection (this was true in 56 studies, 89%). Among them 14 studies (25%) used more than one platform for data collection (for a full overview of studies and data sources, see Appendix A). A single study used custom LMS, two studies used video streaming software, and one study used wiki. Digital traces were the most collected data types in (90.5%), followed by self-reported surveys 27 (42.9%). Self-reported surveys were used to collect data about students' depositions such as engagement, motivation and learning styles. Relational and social network data from computer supported mediated interactions were collected in eight studies (12.7%). Interviews were collected in five studies, video or observation in three, multimodal data were collected in two studies, and, transcripts of classroom interactions were reported in one study.

Most of the data collected in the reviewed studies were digital data (see Table 3). Data were collected from the classroom in only six studies, where two other studies reported on multimodal data, and four studies used video recording and observation of classroom setting. [45] used multimodal data through a system called SPACLE to record classroom interactions among students and teachers. The interactions recorded included on-task, off-task, talking to class, outside or inactivity data. The system allowed for spatial data about positions of the users in the class, and their activity levels. [85] used classroom observations to report on the teachers and students' classroom behaviour, although the methods do not clearly describe in detail what was observed and how it was reported. [81] collected f2f data to measure teaching presence according to the community of inquiry framework. Transcripts of audio recordings of the lessons facilitated the thematic content analysis. Real-time classroom observations were also done. Performance data such as grades or continuous assessment were collected from most studies (88.9%). While LMS data may be informative, it does not capture the f2f learning environment, the process of learning, or the student-teacher or the student-student dynamics. The stark contrast between results collected from digital resources and classroom represents an obvious gap. Most data were gathered using digital traces, dispositional self-reports, relational data and interviews that are disconnected from the classroom where a significant amount of learning happens.

Table 3. Types of data collected and their percentage across all studies

Data type	Y	%	No	%
Trace	57	90.5	6	9.5
Survey	27	42.9	36	57.1
SNA	8	12.7	55	87.3
Interviews	5	7.9	58	92.1
Observation/video	3	4.8	60	95.2
Multimodal	2	3.2	61	96.8
Discourse	2	3.2	61	96.8

The analysis methods in most of the studies (98.4%) employed were the traditional descriptive statistics, frequentists, and group comparisons, that included correlations, comparison of means, and chi-square (see Table 4). Visualisation was used in a significant number of studies (77.8%), in the context of explaining results, but not necessarily as a research objective. Thus, few studies used visualisation as their research objective. However, we also found evidence of development of systems that gather information from different data sources to provide visual analytics to enhance feedback offered to students [102]. However, such application of visualisation was rare.

Table 4. Overview of analysis methods

Methods	Count	%	No	%
Statistics	62	98.4	1	1.6
Visualisation	49	77.8	14	22.2
Regression	32	50.8	31	49.2
Machine learning or AI	29	46.0	34	54.0
Clustering	21	33.3	42	66.7
SNA	10	15.9	53	84.1
Sequence or process mining	10	15.9	53	84.1
Qualitative	9	14.3	54	85.7
Data mining or text analytics	5	7.9	58	92.1

Regression analysis were used to predict performance, or forecast learning outcomes in 29 studies (46%). Results show that prediction of performance is the main research objective for learning analytics in blended learning. In 88.9% of all the studies included performance, prediction or optimisation as the main objective. In 33.3% studies methods for unsupervised classification of students by means of clustering studies were used to categorise students according to certain criteria such as learning strategies, baseline disposition, learning process sequences or self-regulation. Sequence mining appears to be gaining in the learning analytics field with 15.9% of the studies exploring the concept, and, in most of the times it was coupled with clustering and visualisation. Yet, all the studies in this category have not researched the impact of these visualisation on teachers or learners. Studies that used SNA in the analysis are 15.9%, and, similar to process mining research, all of the articles have not used visualisation techniques for the sake of helping students or teacher to optimise learning. Qualitative research was performed in nine studies through the analysis of interviews or transcripts. Data mining and pattern recognition was performed in five studies.

4.3 Ethical and legal considerations

Irrespective of the necessity for considering ethical obligations in the use of student data, the papers did not provide documentations of such responsible use of personal data. Almost all the literature examined in this study, that is 99% of the articles, primarily focuses on LMS data. However, the ethical and legal aspects are very much under-represented in the discussions with only eight articles provided a clear evidence that they do not count on personal data or the data are de-identified. 22 of the 70 reviewed articles (31%) mentioned anonymising students.

Nevertheless, it is important to recall here that hiding the student names from the data set is not enough to guarantee that individuals cannot be identified [40]. For example, if a student who enrolled in a course in a specific year, with specific major and so on could possibly have a significant probability of resulting in a perfect attribute set for identifying a specific student. Such events might raise red flags for ethical concerns of how legible is it to consider that anonymisation of data is sufficient (*ibid.*). Although 40% of the articles indicate that they, at some point, considered ethical aspects when collecting data, which are those ethical aspects were, how did these aspects mattered in the data collection, processing and outcomes, were not been mentioned in any of the reviewed studies. An important observation here is that at least 24 papers among the reviewed studies explicitly focus on the collection, analysis and managing of individual's personal data. Although a more profound discussion to explicate the instrumentation of the legal and ethical procedure of retrieving and processing the sensitive pieces of the data is anticipated, a considerable gap in this focus in the articles is inevitable. Thirteen articles reported studies from Europe, but only six articles are mentioning that they have considered legal aspects and informed the students before the data collection, or the data has been anonymised. As nearly all of the studies were conducted prior to the GDPR rules in the EU [29], new and rigorous practices need to be applied in future LA approaches.

4.4 Contributions

The contributions of reviewed studies could be classified into three themes, such as i) understanding and predicting performance, ii) understand student's behaviours and profiles, and iii) understanding and improving the learning environments (for an overview, see Table 5).

Table 5. Overview of contributions of reviewed studies

Theme	Aim/objectives	Contribution	#
Understand /predict performance	Predict academic performance	Random forest, Linear and logistic regression, ensemble modelling predicts academic performance with satisfying accuracy Forecast model could predict at-risk students Visualisations are helpful for teachers to detect anomalous situations Two to six-week data is enough for future academic performance prediction	11
	Portability of predictive models	Portability of predictive models is low across courses LMS variables vary among general and course-specific models	2
	Association of variables with performance prediction	Data variables related to LMS engagement, self-regulated learning, and collaborative learning are correlated with students' academic performance Tracking data is not significant predictors of academic success for some courses (e.g., graphic design)	9
	Identification of factors affecting learning outcomes	Seven factors found that affect students' academic performance, consisting of four online and three traditional Four factors each found for both collaborative and self-regulated learning that affects students' outcomes	3
	Influence of social network with performance	Social network metrics can be used as predictors Number of interactions do not significantly correlate with student performance SNA based upon questionnaires provide useful indicators for a more fine-grained analysis	3
Understand student's behaviours and profiles	Identification of behaviours /learning patterns	Two learning profiles identified based on student's participation in online activities Identified learning behaviours before and after midterm exams Different self-regulated learning behaviours identified based on resources utilisation and procrastinator nature Found five learning trajectories with the varied resource use	6
	Clustering/profiling-based student learning behaviours	Four student clusters were observed based on their performance measures Based on interactions with the video annotation tool four profiles emerged Six profiles emerged from nine trace variables and student's information system data Based on students' viewing behaviours, they were clustered into three groups Based on usage of LMS three profiles discovered Based on learner control, scaffolding, and interaction three self-regulated profiles appeared	8
	Relation between profiles/learning behaviours on learning outcomes	Self-assessment exercises, regularly resource access, and active online behaviour are significantly correlated with learning outcomes Use of videos annotations, metacognitive skills, and motivational strategies are weakly associated with learning achievements Procrastination behaviour, low level of participation, and worked examples could affect students learning outcome Found that students tend to change their learning behaviour throughout the course	15

Understanding and improving the learning environments	Learning resources and activities in relation to learning outcome	Resources access, LMS access time, and active learning days have a positive influence on learning results Fully worked-out solutions and engagements have adverse effects on students' achievements Personalised feedback has a small to medium positive effect on the learning outcome Visualisations feedback allows students to make a better diagnosis of their performance Learning analytics-based interventions can improve student academic achievement	15
	Improvements in course design, content, and instructions	Video viewing patterns, resource utilisation, and order of activities provide feedback to enhance classroom teaching and resources Visualisation-based learning analytics allow teachers to identify which learning design elements should be revised and improved Differences in instructional approaches during f2f and blended courses are very likely due to the different class formats An understanding develops teachers' interventions through learning activity redesign can cultivate better learning attitudes	8

- 1. Understanding and predicting performance:** To predict students' academic performance, random forest, linear and logistic regression, and ensemble modelling based predictive models provided satisfying results (over 70% accuracy) [2, 51, 77, 80, 81, 102]. Similarly, a forecast learning outcome model (FLOM) was developed using interactive data to predict at-risk students [67]. However, FLOM achieved lower accuracy than other predictive models. On the contrary, student's data visualisations found helpful for teachers to detect anomalous situations [97]. Regarding appropriate time for prediction, studies discovered that two to six weeks data is sufficient to obtain accurate prediction [51, 53]. However, the portability of predictive models across courses remains low [23, 32]. Since prediction is entirely dependent on the supplied data, studies identified that LMS variables (e.g., access time), engagement indicators, self-regulated learning (e.g., self-efficacy and test anxiety), and collaborative learning (e.g., social stability, and time spent on task) variables have reliable predictive power due to their positive correlation with students achievements [2, 30, 71, 80, 90, 91, 102]. Nevertheless, for some courses (e.g., graphic design) tracking data became useless because different patterns exist in the effect of individual data variables [32]. Reviewed studies also disclosed that social network metrics (e.g., degree, authority and PageRank) could be employed to predict student performance [15, 43, 81]. However, using these metrics, the representativity of the predictive models would be limited [81]. In factors identification, four online (e.g., activities, video clicks, videos backwards and practice score per week) and three traditional (e.g., participation in after-school tutoring, homework and quiz scores) factors were identified that affect students' performance [53]. While, attendance, time spent in class, sitting position, and groups are essential for collaborative learning and self-efficacy, positive strategy, less anxious and less usage of negative strategy found important for self-regulated learning [2, 72].
- 2. Understand student's behaviours and profiles:** To identify students learning patterns and behaviours, studies utilised student's participation, resources access, and

other LMS data. For instance, based on students participations two learning behaviours emerged: sensing where students are more likely to participate in information access, interactive and networked learning activities, reflective where students are more predisposed to materials development activities [53]. Similarly, a study identified behaviours before and after midterm exams, for example, out-degree centrality, LMS visit, and time spent before midterm exams and discussion views and visit interval regularity after the midterm exam [51]. In the self-regulated learning context, based on resources access three patterns emerged: self-regulator, external sources users, non-self-regulatory and based on LMS data four behaviours emerged: continuously active, probers, procrastinators and inactive [12, 20]. Likewise, based on resource use five different learning trajectories discovered: overall below-average activity, average resource use, higher use of resources, most active students, least active students [56]. Similarly, studies clustered and profiled students based on their learning behaviours, for example, four clusters (achievers, regular, half-hearted, underachievers) discovered using students' performance measures [57]. Likewise, using video annotation tool interactions, four profiles were created, such as minimalists, task-oriented, disenchanted, and intensive [60]. Correspondingly, students viewing behaviour were adopted to cluster consistent, slide intensive and less intensive students [27]. On the other hand, utilising e-tutorial and information systems data, six profiles emerged, which were the difference in overall activity level and the use of worked-out solutions [62]. In the same vein, based on LMS usages, three clusters were generated such as low, acceptable, and good and students have different patterns of learning behaviour in these clusters [59]. In self-regulated learning context based on authenticity, personalisation, learner control, scaffolding, and interaction, three profiles were identified such as self-regulating, external regulating and lack of regulation [21]. On the other hand, a considerable number of studies contributed in terms of identifying the association and effects of different learning behaviour on students' achievements. For instance, self-assessment exercises, regularly resources access, active online behaviour, and time management are significantly correlated with student learning outcome [5, 18, 23, 34, 49, 52, 63, 79]. While, the use of videos annotations, metacognitive skills, and motivational strategies are weakly associated with learning achievements [54, 55, 73]. On the other side, procrastination behaviour, low level of participation, and dependency on worked examples could affect students learning outcome [54, 60, 89]. Furthermore, few studies discovered that students have a tendency to change their learning behaviour throughout the course and comparison can be conducted between successful and non-successful students based on their learning patterns [34, 49].

- 3. Understanding and improving the learning environments:** Reviewed studies discovered that course material access without lapses, LMS access time, active learning days and teachers' monitoring influence learning results [1, 7, 8, 44, 45, 65]. Whereas, worked-out solutions and engagements create adverse effects on students' achievements [41, 66, 85]. In the context of feedback provision, personalised feedback have a small to medium positive effect on the learning outcome [71]. In terms of intervention, learning analytics-based interventions improved student academic achievement, with a 10.6% higher score than blended learning without intervention

[36]. Concerning improvement in courses, video viewing patterns, resource utilisation, course item frequencies and order of activities provide enough feedback to enhance classroom teaching and course resources [19, 23, 34]. Similarly, visualisation-based learning analytics allow teachers to identify which learning design elements should be revised and improved [59].

5 Discussion

RQ1 and RQ2: We raised the questions how blended learning is defined and how learning theories and perspectives are used in the reviewed learning analytics research?

In line with [46] we conclude that BL seems to have become somewhat of a meta-concept. Thus, as have been detailed in the results, blended learning is often not an adoption of *one pure* type of blended learning, but a combination of blended learning approaches. When BL approaches are combined, there is a greater complexity than in a "simple blend", why we propose that this can be referred to as *complex hybrid learning*. For example, we see combinations of blended instruction *and* blended distribution: *optional adoption* [e.g. 66, 80], *synchronous teaching of f2f students and online students* [6] and iii) *reversed distribution*; [35, 40]. Results show that SRL, by far, was the most common used theory. To operationalise SRL, engagement and other perspectives of learning, LMS trace data was used to collate the number of messages sent, documents uploaded, chat sessions attended, as well as data collected to analyse teacher presence (e.g. 1, 54, 79). In line with [42], we conclude, that operationalisations relying on LMS data might risk to be superficial and oversimplified interpretations of the underpinning theory. However, some of the reviewed studies also explored relations between trace data, performance and SRL using self-reports [70, 87, 91]. However, results revealed that certain perspectives of learning were more common to explore than others: for example, the flipped classroom, collaborative learning and conversational aspects of learning. We thus call for innovative perspectives of learning, for example, complex or multi-modal data gathering, longitudinal studies and mixed methods approaches. Conclusively, the theoretical underpinnings of a research study (including what is meant by BL and the learning perspectives taken), are needed to increase clarity, quality and validity of the objectives and contributions of that study to enable comparison and transparency a richer description of the actual blended learning environment approached.

RQ 3: We also raised the question, what approaches of data collection, methods and analysis that are evident in the reviewed learning analytics research. One would expect that learning analytics in a blended learning scenario would account for the fact that the context of BL integrates both modalities (physical and digital). However, most of the reviewed studies have used digital traces, dispositional self-reports, relational data and interviews that does not fully cover the gamut of possible data sources of the classroom where a significant amount of learning happens. While predictive models have -in many cases- been able to infer future performance, they have failed short of explaining learning, or offer a guide on how to intervene in the classroom. Of course, collecting data in a blended scenario is not easy, and therefore more research is needed that collects contextually relevant data, and more importantly, on how to unobtrusively collect data in

the classroom. We also recognise the benefits of visualisation as an intermediate step: albeit we found that visual analytics were rare [102]. Believing that eye-balling the data, the accessibility of instant overview, might support the teacher, we propose that more research on the impact of visual insights offered to stakeholders is needed. We found that blended learning studies did not use classroom data to investigate the complexity of both the online and f2f learning setting.

RQ4: We then explored how ethical and legal aspects are considered in the reviewed learning analytics research. In line with previous critique [e.g. 74; 84] we found that although ethical, privacy and legislator requirements exist, the current practises do not always consider these. While results reveal that almost all (99%) of the reviewed studies were conducted prior to the GDPR rules in the EU. 13 articles reported studies from Europe, but only six articles mentioned legal aspects, having informed the students prior to data collection, or considered anonymising the data. This raises critical concerns, as aside the GDPR legislation, ethical considerations need still be adhered to. This may also reflect general slow governmental responses to regulate consequences of the digitalisation.

RQ5: Lastly, we surveyed the contributions of the reviewed articles. The results revealed that reviewed articles made several contributions on predicting academic performance, identifying learning behaviours, and improving learning environments. With regards to predicting academic performance, machine learning-based predictive models was proven to be effective but with low portability across courses, whereas visualisation-based methods required teacher assistance [2, 77]. Moreover, data variables effectiveness on performance prediction is based on course structure; however, social network metrics and variables related to LMS engagement, self-regulated learning, collaborative learning are found significantly correlated with academic performance [86, 90, 102]. In terms of identifying learning behaviours, results show that by utilising student's participation in online activities and resources access impactful learning behaviours could be identified, and these behaviours are beneficial to cluster or profile students based on adopted behaviours [56]. In the regards of improving learning environments, results show that learning resources that provide student assistance to complete their assignments create positive effects on learning outcome [1, 7].

6 Conclusion and Future Research

As BL currently seems to be a more general concept, detailed descriptions of the actual learning situation, delivery, blend or hybrid solution is needed alongside clear underpinning theories to position the research, or as proposed an indication of whether one is approaching a "simple blend" or complex hybrid learning. We argue that in the current wake of the transforming distance learning, we see hybrid solutions, that raises awareness of a complexity of multiple blended solutions in parallel, that if not described, could mean just about any kind of learning, delivery or setting. We found that data used in many learning analytics studies were used as a proxy for what happens in the classroom. However, when studies do not include manifestations in the real classroom, they fall short of explaining learning, or offer a guide on how to intervene in the

classroom. More research is needed that accounts for the context of BL and more importantly, on how to unobtrusively collect relevant data that enables the support of learning where it occurs. In the light of our findings of ethical and legal considerations, we strongly argue that while there are no established traditions in LA research in terms of legal requirements; new and rigorous practices need to be developed and applied in current and future LA approaches. Ethical consequences might be devastating and the field urgently needs to acknowledge this lack of consideration.

7 Acknowledgement

This research is financed by the Swedish Research Council (VR).

8 References

- [1] Agnihotri, L., Essa, A., & Baker, R. (2017). Impact of student choice of content adoption delay on course outcomes. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (pp. 16–20). ACM. <https://doi.org/10.1145/3027385.3027437>
- [2] Akhtar, S., Warburton, S., & Xu, W. (2017). The use of an online learning and teaching system for monitoring computer aided design student participation and predicting student success. *International Journal of Technology and Design Education*, 27(2), 251–270. <https://doi.org/10.1007/s10798-015-9346-8>
- [3] Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37(April 2018), 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- [4] Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). Applications of data science to game learning analytics data: A systematic literature review. *Computers and Education*, 141(June), 103612. <https://doi.org/10.1016/j.compedu.2019.103612>
- [5] Andergassen, M., Mödritscher, F., & Neumann, G. (2014). Practice and Repetition during Exam Preparation in Blended Learning Courses: Correlations with Learning Results. *Journal of Learning Analytics*, 1(1), 48–74. <https://doi.org/10.18608/jla.2014.11.4>
- [6] Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning Analytics in Distance Education: A Systematic Literature Review. *Online Learning*, 20(2 (October)), 13–29. <https://doi.org/10.24059/olj.v20i2.790>
- [7] Ayub, M., Toba, H., Yong, S., & Wijanto, M. C. (2017). Modelling students' activities in programming subjects through educational data mining. *Global Journal of Engineering Education*, 19(3), 249–255. <https://doi.org/10.1109/icodse.2017.8285881>
- [8] Baik, E. J., & Reynolds, R. B. (2013). Contribution of wiki trace and wiki resource use variables towards quality of game design in a guided discovery-based program of game design learning. *Proceedings of the ASIST Annual Meeting*, 50(1). <https://doi.org/10.1002/meet.14505001165>
- [9] Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior*, 71, 550–585. <https://doi.org/10.1016/j.chb.2016.08.049>

- [10] Bodily, R., Kay, J., Alevan, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). Open Learner Models and Learning Analytics Dashboards: A Systematic Review in Proceedings of the 8th international conference on learning analytics and knowledge (pp. 41-50). <https://doi.org/10.1145/3170358.3170409>
- [11] Boelens, R., Van Laer, S., De Wever, B., & Elen, J. (2015). Blended learning in adult education: towards a definition of blended learning. *Adult Learners Online! Blended and Online Learning in Adult Education and Training*. <https://doi.org/10.21125/iceri.2018.1219>
- [12] Bos, N., & Saskia, B. G. (2016). Student differences in regulation strategies and their use of learning resources: Implications for educational design. In *ACM International Conference Proceeding Series* (Vol. 25-29-April, pp. 344–353). New York, New York, USA: Association for Computing Machinery. [doi:10.1145/2883851.2883890](https://doi.org/10.1145/2883851.2883890)
- [13] Bray, N. J., & Major, C. H. (2012). Status of Journals in the Field of Higher Education. *The Journal of Higher Education*, 82(4), 479–503. <https://doi.org/10.1080/00221546.2011.11777213>
- [14] Cerezo, R., Esteban, M., Sánchez-Santillán, M., & Núñez, J. C. (2017). Procrastinating behavior in computer-based learning environments to predict performance: A case study in Moodle. *Frontiers in Psychology*, 8, 1–11. <https://doi.org/10.3389/fpsyg.2017.01403>
- [15] Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 1165-1188. <https://doi.org/10.2307/41703503>
- [16] Chen, B., Knight, S., & Wise, A. (2018). Critical issues in designing and implementing temporal analytics. *Journal of Learning Analytics*, 5 (1) (2018), p. 9. <https://doi.org/10.18608/jla.2018.51.1>
- [17] Chen, J., & Fong, D. (2020). A Motivational Story in Hong Kong: Generating Goals for Language Learners and Blended Learning Designers from a Mixed-Method Learning Analytics Approach in English for Academic Purposes. In *Technology and the Psychology of Second Language Learners and Users* (pp. 491-516). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-34212-8_19
- [18] Cheng, G., & Chau, J. (2016). Exploring the relationships between learning styles, online participation, learning achievement and course satisfaction: An empirical study of a blended learning course. *British Journal of Educational Technology*, 47(2), 257–278. <https://doi.org/10.1111/bjet.12243>
- [19] Chetlur, M., Tamhane, A., Reddy, V. K., Sengupta, B., Jain, M., Sukjunnimit, P., & Wagh, R. (2014). EduPaL: Enabling Blended Learning in Resource Constrained Environments. In *Proceedings of the Fifth ACM Symposium on Computing for Development* (pp. 73–82). ACM. <https://doi.org/10.1145/2674377.2674388>
- [20] Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessel, A., Barreiros, C., & Lindstädt, S. (2018). Finding traces of self-regulated learning in activity streams. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 191–200). ACM. <https://doi.org/10.1145/3170358.3170381>
- [21] Clow, D. (2012a). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 134-138). doi.org/10.1145/2330601.2330636
- [22] Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012b). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, 26(17), 2583-2591. <https://doi.org/10.1002/hyp.9385>
- [23] Conijn, R., Van den Beemt, A., & Cuijpers, P. (2018). Predicting student performance in a blended MOOC. *Journal of Computer Assisted Learning*, 34(5), 615–628. <https://doi.org/10.1111/jcal.12270>

- [24] Conijn, Rianne, Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using moodle LMS. *IEEE Transactions on Learning Technologies*, 10(1), 17–29. <https://doi.org/10.1109/tlt.2016.2616312>
- [25] Dawson, S., Mirriahi, N., & Gasevic, D. (2015). Importance of theory in learning analytics in formal and workplace settings. *Journal of Learning Analytics*, 2(2), 1-4. <https://doi.org/10.18608/jla.2015.22.1>
- [26] De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic bulletin & review*, 20(4), 631-642. <https://doi.org/10.3758/s13423-013-0386-3>
- [27] De-Marcos, L., García-López, E., García-Cabot, A., Medina-Merodio, J. A., Domínguez, A., Martínez-Herraíz, J. J., & Díez-Follo, T. (2016). Social network analysis of a gamified e-learning course: Small-world phenomenon and network metrics as predictors of academic performance. *Computers in Human Behavior*, 60, 312–321. <https://doi.org/10.1016/j.chb.2016.02.052>
- [28] Dimić, G., Predić, B., Rančić, D., Petrović, V., Maček, N., & Spalević, P. (2018). Association analysis of moodle e-tests in blended learning educational environment. *Computer Applications in Engineering Education*, 26(3), 417–430. <https://doi.org/10.1002/cae.21894>
- [29] EU 2016/679. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. Brussels: European Commission. <https://doi.org/10.5593/sgemsocial2019v1.1/s02.022>
- [30] Gamulin, J., Gamulin, O., & Kermek, D. (2016). Using Fourier coefficients in time series analysis for student performance prediction in blended learning environments. *Expert Systems*, 33(2), 189–200. <https://doi.org/10.1111/exsy.12142>
- [31] Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *Internet and Higher Education*, 7(2), 95–105. <https://doi.org/10.1016/j.iheduc.2004.02.001>
- [32] Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *Internet and Higher Education*, 28, 68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>
- [33] Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. *Journal of Learning Analytics*, 4(2), 113–128. <https://doi.org/10.18608/jla.2017.42.10>
- [34] Gelan, A., Fastré, G., Verjans, M., Martin, N., Janssenswillen, G., Creemers, M., ... Thomas, M. (2018). Affordances and limitations of learning analytics for computer-assisted language learning: a case study of the VITAL project. *Computer Assisted Language Learning*, 31(3), 294–319. <https://doi.org/10.1080/09588221.2017.1418382>
- [35] Gewerc-Barujel, A., Montero-Mesa, L., & Lama-Peñín, M. (2013). Collaboration and Social Networking in Higher Education. *Comunicar*, 21(42), 55–63. <https://doi.org/10.3916/c42-2014-05>
- [36] Gong, L., Liu, Y., & Zhao, W. (2018). Using Learning Analytics to Promote Student Engagement and Achievement in Blended Learning: An Empirical Study. In *Proceedings of the 2nd International Conference on E-Education, E-Business and E-Technology*. <https://doi.org/10.1145/3241748.3241760>

- [37] Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In B. Miller (Ed.), *The handbook of blended learning* (1st ed., Vol. LVII, pp. 3–21). San Francisco, CA: Pfeiffer.
- [38] Graham, C.R. (2009) Blended Learning Models. In: *Encyclopedia of Information Science and Technology*. Hershey, PA: Idea Group Inc., 375-383.
- [39] Graham, C. R. (2013). Emerging practice and research in blended learning. In M. G. Moore (Ed.), *Handbook of Distance Education* (3rd ed., pp. 333–350). New York, New York, USA: Routledge.
- [40] Harrak, F., Bouchet, F., Luengo, V., & Gillois, P. (2018). Profiling students from their questions in a blended learning environment. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 102–110). ACM. <https://doi.org/10.1145/3170358.3170389>
- [41] Hecking, T., Ziebarth, S., & Hoppe, H. U. (2014). Analysis of Dynamic Resource Access Patterns in Online Courses. *Journal of Learning Analytics*, 1(3), 34–60. <https://doi.org/10.18608/jla.2014.13.4>
- [42] Henrie, C. R., Bodily, R., Larsen, R., & Graham, C. R. (2018). Exploring the potential of LMS log data as a proxy measure of student engagement. *Journal of Computing in Higher Education*, 30(2), 344–362. <https://doi.org/10.1007/s12528-017-9161-1>
- [43] Hernández-Nanclares, N., García-Muñiz, A. S., & Rienties, B. (2017). Making the most of “external” group members in blended and online environments. *Interactive Learning Environments*, 25(4), 467–481. <https://doi.org/10.1080/10494820.2016.1140656>
- [44] Hill, T., Chidambaram, L., & Summers, J. D. (2017). Playing ‘catch up’ with blended learning: performance impacts of augmenting classroom instruction with online learning. *Behaviour and Information Technology*, 36(1), 54–62. <https://doi.org/10.1080/0144929x.2016.1189964>
- [45] Holstein, K., McLaren, B. M., & Aleven, V. (2017). SPACLE: investigating learning across virtual and physical spaces using spatial replays. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, 358–367. <https://doi.org/10.1145/3027385.3027450>
- [46] Hrastinski, S. (2019). What Do We Mean by Blended Learning? *TechTrends*, 63(5), 564–569. <https://doi.org/10.1007/s11528-019-00375-5>
- [47] Hui, Y. K., Mai, B., Qian, S., & Kwok, L. F. (2018). Cultivating better learning attitudes: a preliminary longitudinal study. *Open Learning*, 33(2), 155–170. <https://doi.org/10.1080/02680513.2018.1454830>
- [48] Johnson, L., Becker, S. A., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC Horizon Report; 2016 Higher Education Edition*. Retr. Feb. 13, 2019, from www.sconul.ac.uk/sites/default/files/documents/2016-nmc-horizon-report-he-EN-1.pdf
- [49] Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *Internet and Higher Education*, 33, 74–85. <https://doi.org/10.1016/j.iheduc.2017.02.001>
- [50] Khalil, M., & Ebner, M. (2016). De-identification in learning analytics. *Journal of Learning Analytics*. <https://doi.org/10.18608/jla.2016.31.8>
- [51] Kim, D., Park, Y., Yoon, M., & Jo, I. H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *Internet and Higher Education*, 30, 30–43. <https://doi.org/10.1016/j.iheduc.2016.03.002>
- [52] Li, L. Y., & Tsai, C. C. (2017). Accessing online learning material: Quantitative behavior patterns and their effects on motivation and learning performance. *Computers and Education*, 114(300), 286–297. <https://doi.org/10.1016/j.compedu.2017.07.007>

- [53] Lu, O. H. T., Huang, A. Y. Q., Lin, A. J. Q., Ogata, H., & Yang, S. J. H. (2018). Applying Learning Analytics for the Early Prediction of Students' Academic Performance in Blended Learning. *Educational Technology & Society* (Vol. 21).
- [54] Lukarov, V., Verbert, K., & Schroeder, U. (2019). Scaling up learning analytics in blended learning scenarios (Doctoral dissertation, Universitätsbibliothek der RWTH Aachen).
- [55] Manzanares, M. C. S., Sánchez, R. M., García Osorio, C. I., & Díez-Pastor, J. F. (2017). How do B-learning and learning patterns influence learning outcomes? *Frontiers in Psychology*, 8(MAY), 1–13. <https://doi.org/10.3389/fpsyg.2017.00745>
- [56] Marsh, J., & Drexler, P. (2001). How to design effective blended learning. Sunnyvale: Brandon-Hall. Sunnyvale: California, USA.
- [57] Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2019). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, 1382(c), 1–1. <https://doi.org/10.1109/tlt.2019.2916802>
- [58] McKenzie, W. A., Perini, E., Rohlf, V., Toukhsati, S., Conduit, R., & Sanson, G. (2013). A blended learning lecture delivery model for large and diverse undergraduate cohorts. *Computers and Education*, 64, 116–126. <https://doi.org/10.1016/j.compedu.2013.01.009>
- [59] Melero, J., Hernández-Leo, D., Sun, J., Santos, P., & Blat, J. (2015). How was the activity? A visualization support for a case of location-based learning design. *British Journal of Educational Technology*, 46(2), 317–329. <https://doi.org/10.1111/bjjet.12238>
- [60] 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(6), 1083–1106. <https://doi.org/10.1007/s11423-016-9449-2>
- [61] Misiejuk, K., & Wasson, B. (2017). State of the Field report on Learning Analytics. SLATE Report 2017-2.
- [62] Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, 9(2), 75–85. <https://doi.org/10.1007/s11409-014-9114-2>
- [63] Montgomery, A. P., Mousavi, A., Carbonaro, M., Hayward, D. V., & Dunn, W. (2019). Using learning analytics to explore self-regulated learning in flipped blended learning music teacher education. *British Journal of Educational Technology*, 50(1), 114–127. <https://doi.org/10.1111/bjjet.12590>
- [64] Musabirov, I., & Bakhitova, A. (2017). Trajectories of student interaction with learning resources in blended learning. In *Proceedings of the 17th Koli Calling International Conference on Computing Education Research* (pp. 191–192). ACM. <https://doi.org/10.1145/3141880.3141907>
- [65] Mödritscher, F., Andergassen, M., & Neumann, G. (2013). Dependencies between E-Learning Usage Patterns and Learning Results. In *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies* (pp. 1–8). ACM. <https://doi.org/10.1145/2494188.2494206>
- [66] Nguyen, Q., Tempelaar, D. T., Rienties, B., & Bas, G. (2016). What learning analytics-based prediction models tell us about feedback preferences of students. *Q. Rev. Distance Education*, 17(3), 13–33.
- [67] Nguyen, V. A., Nguyen, Q. B., & Nguyen, V. T. (2018). A Model to Forecast Learning Outcomes for Students in Blended Learning Courses Based on Learning Analytics, 35–41. <https://doi.org/10.1145/3268808.3268827>
- [68] Nouri, J., Ebner, M., Ifenthaler, D., Saqr, M., Malmberg, J., Khalil, M., ... Berthelsen, U. D. (2019). Efforts in Europe for Data-Driven Improvement of Education – A Review of

- Learning Analytics Research in Seven Countries. *International Journal of Learning Analytics and Artificial Intelligence for Education (IJAI)*, 1(1), 8. <https://doi.org/10.3991/ijai.v1i1.11053>
- [69] Nouri, J., Saqr, M & Fors, U. (2019). Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback. *Journal of Systemics, Cybernetics and Informatics*. 17(2).
- [70] Pardo, A., Han, F., & Ellis, R. A. (2016). Exploring the relation between self-regulation, online activities, and academic performance, 422–429. doi:10.1145/2883851.2883883
- [71] Pardo, A., Han, F., & Ellis, R. A. (2017). Combining University student self-regulated learning indicators and engagement with online learning events to Predict Academic Performance. *IEEE Transactions on Learning Technologies*, 10(1), 82–92. <https://doi.org/10.1109/tlt.2016.2639508>
- [72] Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, 50(1), 128–138. <https://doi.org/10.1111/bjet.12592>
- [73] Pardo, A., Mirriahi, N., Dawson, S., Zhao, Y., Zhao, A., & Gašević, D. (2015). Identifying learning strategies associated with active use of video annotation software, 255–259. <https://doi.org/10.1145/2723576.2723611>
- [74] Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438–450. <https://doi.org/10.1111/bjet.12152>
- [75] Park, Y., Yu, J. H., & Jo, I. H. (2016). Clustering blended learning courses by online behavior data case study in a Korean higher education institute. *Internet and Higher Education*, 29, 1–11. <https://doi.org/10.1016/j.iheduc.2015.11.001>
- [76] Paskevicius, M., & Bortolin, K. (2016). Blending our practice: using online and face-to-face methods to sustain community among faculty in an extended length professional development program. *Innovations in Education and Teaching International*, 53(6), 605–615. <https://doi.org/10.1080/14703297.2015.1095646>
- [77] Predić, B., Dimić, G., Rančić, D., Štrbac, P., Maček, N., & Spalević, P. (2018). Improving final grade prediction accuracy in blended learning environment using voting ensembles. *Computer Applications in Engineering Education*, 26(6), 2294–2306. <https://doi.org/10.1002/cae.22042>
- [78] Reeve, J. (2012). A Self-determination Theory Perspective on Student Engagement. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 149–172). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4614-2018-7_7
- [79] Rubio, F., Thomas, J. M., & Li, Q. (2018). The role of teaching presence and student participation in Spanish blended courses. *Computer Assisted Language Learning*, 31(3), 226–250. <https://doi.org/10.1080/09588221.2017.1372481>
- [80] Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical Teacher*, 39(7), 757–767. <https://doi.org/10.1080/0142159x.2017.1309376>
- [81] Saqr, M., Fors, U., & Tedre, M. (2018). How the study of online collaborative learning can guide teachers and predict students' performance in a medical course. *BMC Medical Education*, 18(1), 1–14. <https://doi.org/10.1186/s12909-018-1126-1>
- [82] Scholes, V. (2016). The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, 64(5), 939–955. <https://doi.org/10.1007/s11423-016-9458-1>
- [83] Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>

- [84] Slade, S., & Prinsloo, P. (2013). Learning Analytics. *American Behavioral Scientist*, 57(10), 1510–1529. <https://doi.org/10.1177/0002764213479366>
- [85] Snodgrass Rangel, V., Bell, E. R., Monroy, C., & Whitaker, J. R. (2015). Toward a New Approach to the Evaluation of a Digital Curriculum Using Learning Analytics. *Journal of Research on Technology in Education*, 47(2), 89–104. <https://doi.org/10.1080/15391523.2015.999639>
- [86] Tempelaar, D. (2017). How Dispositional Learning Analytics Helps Understanding the Worked-Example Principle. *International Association for Development of the Information Society*, (Celda), 14.
- [87] Tempelaar, D., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018a). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior*, 78, 408–420. <https://doi.org/10.1016/j.chb.2017.08.010>
- [88] Tempelaar, D., Rienties, B., & Nguyen, Q. (2018b). A multi-modal study into students' timing and learning regulation: time is ticking. *Interactive Technology and Smart Education*, 15(4), 298–313. <https://doi.org/10.1108/itse-02-2018-0015>
- [89] Tempelaar, D., Rienties, B., & Nguyen, Q. (2018c). Investigating learning strategies in a dispositional learning analytics context, 201–205. <https://doi.org/10.1145/3170358.3170385>
- [90] Tempelaar, D., 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. <https://doi.org/10.1016/j.chb.2014.05.038>
- [91] Tempelaar, D., Rienties, B., Nguyen, Q., Macan, T., & Hoffimacan, T. (2017). Towards Actionable Learning Analytics Using Dispositions, 79(1), 381–391. doi:10.1037/0021-9010.79.3.381
- [92] Tempelaar, D. (2020). Supporting the less-adaptive student: the role of learning analytics, formative assessment and blended learning. *Assessment & Evaluation in Higher Education*, 45(4), 579-593. <https://doi.org/10.1080/02602938.2019.1677855>
- [93] van Goidsenhoven, S., Bogdanova, D., Deeva, G., Broucke, S. V., De Weerd, J., & Snoeck, M. (2020, March). Predicting student success in a blended learning environment. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 17-25). <https://doi.org/10.1145/3375462.3375494>
- [94] Van Laer, S., & Elen, J. (2018). Adults' Self-Regulatory Behaviour Profiles in Blended Learning Environments and Their Implications for Design. *Technology, Knowledge and Learning*, 1–31. <https://doi.org/10.1007/s10758-017-9351-y>
- [95] Van Leeuwen, A. (2019). Teachers' perceptions of the usability of learning analytics reports in a flipped university course: when and how does information become actionable knowledge? *Educational Technology Research and Development*, 67(5), 1043-1064. <https://doi.org/10.1007/s11423-018-09639-y>
- [96] Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers and Education*, 122(March), 119–135. <https://doi.org/10.1016/j.compedu.2018.03.018>
- [97] Villamañe, M., Álvarez, A., Larrañaga, M., Hernández-Rivas, O., & Caballero, J. (2018). Using Visualizations to Improve Assessment in Blended Learning Environments, 165–169. <https://doi.org/10.1145/3284179.3284209>
- [98] Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In *Metacognition in educational theory and practice*.
- [99] Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In *Handbook of self-regulation* (pp. 531-566). Academic Press. <https://doi.org/10.1016/b978-012109890-2/50045-7>

- [100] Wise AF, Shaffer DW. 2015. Why theory matters more than ever in the age of big data. *JLA*. 2:5–13.
- [101] Whitelock-Wainwright, A., Tsai, Y. S., Lyons, K., Kaliff, S., Bryant, M., Ryan, K., & Gašević, D. (2020, March). Disciplinary differences in blended learning design: a network analytic study. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 579-588). <https://doi.org/10.1145/3375462.3375488>
- [102] Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *Internet and Higher Education*, 27, 44–53. <https://doi.org/10.1016/j.iheduc.2015.05.002>

9 Authors

Nina Bergdahl is a researcher in didactics and digitalisation at the Department of Adult and Upper Secondary Education, Malmö, Sweden and affiliated with the Department of Computer and Systems Sciences (DSV) Stockholm university, Sweden. Email: ninabe@dsv.su.se

Jalal Nouri is an associate professor at the Department of Computer and Systems Sciences (DSV) Stockholm university, Sweden.

Thashmee Karunaratne is an associate professor at the Department of Computer and Systems Sciences (DSV) Stockholm university, Sweden

Muhammad Afzaal is a PhD student at the Department of Computer and Systems Sciences (DSV) Stockholm university, Sweden.

Mohammad Saqr is a postdoc at KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, Stockholm, Sweden.

Article submitted 2020-08-19. Resubmitted 2020-09-16. Final acceptance 2020-09-16. Final version published as submitted by the authors.

10 Appendix A

S	Study	LMS	Trace	MML	Observation/video	Interviews	Discourse	SN A	Survey	Classroom	Performance
1	Agnihotri, L., Essa, A., & Baker, R. (2017). Impact of student choice of content adoption delay on course outcomes. In <i>Proceedings of the Seventh International Learning Analytics & Knowledge Conference</i> (pp. 16–20). ACM. https://doi.org/10.1145/3027385.3027437	Y	Y	N	N	N	N	N	N	N	Y
2	Akhtar, S., Warburton, S., & Xu, W. (2017). The use of an online learning and teaching system for monitoring computer aided design student participation and predicting student success. <i>International Journal of Technology and Design Education</i> , 27(2), 251–270. https://doi.org/10.1007/s10798-015-9346-8	Y	Y	Y	N	Y	N	Y	N	Y	Y
3	Andergassen, M., & Mödritscher, F. (2014). Practice and repetition during exam preparation in blended learning courses: Correlations with learning results. <i>Of Learning Analytics</i> , 1, 48–74.	M	Y	N	N	N	N	N	D	N	Y
4	Ayub, M., Toba, H., Yong, S., & Wijanto, M. C. (2017). Modelling students' activities in programming subjects through educational data mining. <i>Global Journal of Engineering Education</i> , 19(3)	Y	Y	N	N	N	N	N	N	N	Y
5	Baik, E. J., & Reynolds, R. B. (2013). Contribution of wiki trace and wiki resource use variables towards quality of game design in a guided discovery-based program of game design learning. <i>Proceedings of the ASIST Annual Meeting</i> , 50(1). https://doi.org/10.1002/meet.14505001165	Wiki	Y	N	N	N	N	N	Y	N	Y
6	Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. <i>Computers in Human Behavior</i> , 71, 550–585. https://doi.org/10.1016/j.chb.2016.08.049	Y	Y	N	N	N	N	N	D	N	Y
7	Bos, N. (n.d.). Student differences in regulation strategies and their use of learning resources: implications for educational design. <i>Learning Analytics and Knowledge (LAK'16)</i> , 27-29 April 2016, 344–353.	Y	Y	N	N	N	N	N	D	Y	Y
8	Cheng, G., & Chau, J. (2016). Exploring the relationships between learning styles, online participation, learning achievement and course satisfaction: An empirical study of a blended learning course. <i>British Journal of Educational Technology</i> , 47(2), 257–278. https://doi.org/10.1111/bjet.12243	Y	Y	N	N	N	N	N	D	N	Y

9	Chen, J., & Fong, D. (2020). A Motivational Story in Hong Kong: Generating Goals for Language Learners and Blended Learning Designers from a Mixed-Method Learning Analytics Approach in English for Academic Purposes. In <i>Technology and the Psychology of Second Language Learners and Users</i> (pp. 491-516). Palgrave Macmillan, Cham.	Y	Y	N	N	N	N	N	D	N	Y
10	Chetlur, M., Tamhane, A., Reddy, V. K., Sengupta, B., Jain, M., Sukjunnimit, P., & Wagh, R. (2014). EduPaL: Enabling Blended Learning in Resource Constrained Environments. In <i>Proceedings of the Fifth ACM Symposium on Computing for Development</i> (pp. 73–82). ACM. https://doi.org/10.1145/2674377.2674388	M	Y	N	N	N	N	N	D	N	Y
11	Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessler, A., Barreiros, C., & Lindstädt, S. (2018). Finding traces of self-regulated learning in activity streams. In <i>Proceedings of the 8th International Conference on Learning Analytics and Knowledge</i> (pp. 191–200). ACM. https://doi.org/10.1145/3170358.3170381	Custom LMS	Y	N	N	N	N	N	Y	N	Y
12	Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data. <i>IEEE Transactions on Learning Technologies</i> , 10(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312	Y	Y	N	N	N	N	N	N	N	Y
13	Conijn, R., Van den Beemt, A., & Cuijpers, P. (2018). Predicting student performance in a blended MOOC. <i>Journal of Computer Assisted Learning</i> , 34(5), 615–628. https://doi.org/10.1111/jcal.12270	M	Y	N	N	N	N	N	N	N	Y
14	Dimić, G., Predić, B., Rančić, D., Petrović, V., Maček, N., & Spalević, P. (2018). Association analysis of moodle e-tests in blended learning educational environment. <i>Computer Applications in Engineering Education</i> , 26(3), 417–430. https://doi.org/10.1002/cae.21894	Y	Y	N	N	N	N	N	N	N	Y
15	De-Marcos, L., García-López, E., García-Cabot, A., Medina-Merodio, J. A., Domínguez, A., Martínez-Herraiz, J. J., & Díez-Folledo, T. (2016). Social network analysis of a gamified e-learning course: Small-world phenomenon and network metrics as predictors of academic performance. <i>Computers in Human Behavior</i> , 60, 312–321. https://doi.org/10.1016/j.chb.2016.02.052	M	N	N	N	N	N	Y	N	N	Y
16	Dobashi, K. (2016). Development and Trial of Excel Macros for Time Series Cross Section Monitoring of Student	Y	Y	N	N	N	N	N	N	N	N

	Engagement: Analyzing Students' Page Views of Course Materials. <i>Procedia Computer Science</i> , 96, 1086–1095. https://doi.org/10.1016/j.procs.2016.08.133										
17	Gamulin, J., Gamulin, O., & Kermek, D. (2016). Using Fourier coefficients in time series analysis for student performance prediction in blended learning environments. <i>Expert Systems</i> , 33(2), 189–200. https://doi.org/10.1111/exsy.12142	Y	Y	N	N	N	N	N	N	N	Y
18	Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. <i>Internet and Higher Education</i> , 28, 68–84. https://doi.org/10.1016/j.iheduc.2015.10.002	Y	Y	N	N	N	N	N	N	N	Y
19	Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. <i>Journal of Learning Analytics</i> , 4(2), 113–128. https://doi.org/10.18608/jla.2017.42.10	Y	Y	N	N	N	N	N	D	N	Y
20	Gelan, A., Fastré, G., Verjans, M., Martin, N., Janssenswillen, G., Creemers, M., ... Thomas, M. (2018). Affordances and limitations of learning analytics for computer-assisted language learning: a case study of the VITAL project. <i>Computer Assisted Language Learning</i> , 31(3), 294–319. https://doi.org/10.1080/09588221.2017.1418382	Y	Y	N	N	N	N	N	N	N	Y
21	Gewerc-Barujel, A., Montero-Mesa, L., & Lama-Penin, M. (2013). Collaboration and Social Networking in Higher Education. <i>Comunicar</i> , 21(42), 55–63. https://doi.org/10.3916/c42-2014-05	Y	Y	N	N	N	N	N	N	N	N
22	Gong, L., Liu, Y., & Zhao, W. (2018). Using Learning Analytics to Promote Student Engagement and Achievement in Blended Learning: An Empirical Study. In <i>Proceedings of the 2nd International Conference on E-Education, E-Business and E-Technology</i> . https://doi.org/10.1145/3241748.3241760	Y	N	N	N	N	N	N	D	N	Y
23	Harrak, F., Bouchet, F., Luengo, V., & Gillois, P. (2018). Profiling students from their questions in a blended learning environment. In <i>Proceedings of the 8th International Conference on Learning Analytics and Knowledge</i> (pp. 102–110). ACM. https://doi.org/10.1145/3170358.3170389	Y	N	N	N	N	N	N	N	Y	N
24	Hecking, T., Ziebarth, S., & Hoppe, H. U. (2014). Analysis of Dynamic Resource Access Patterns in Online Courses. <i>Journal of Learning Analytics</i> ,	Y	Y	N	N	N	N	Y	N	N	N

	1(3), 34–60. https://doi.org/10.18608/jla.2014.13.4										
25	Hernández-Nanclares, N., García-Muñiz, A. S., & Rienties, B. (2017). Making the most of “external” group members in blended and online environments. <i>Interactive Learning Environments</i> , 25(4), 467–481. https://doi.org/10.1080/10494820.2016.1140656	Y	N	N	N	N	N	Y	N	N	N
26	Hill, T., Chidambaram, L., & Summers, J. D. (2017). Playing ‘catch up’ with blended learning: performance impacts of augmenting...: Discovery Service for De La Salle Univ. <i>Behaviour & Information Technology</i> , 36(1), 54–62.	Y	Y	N	N	N	N	N	N	N	Y
27	Holstein, K., McLaren, B. M., & Aleven, V. (2017). SPACLE: investigating learning across virtual and physical spaces using spatial replays. <i>Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17</i> , 358–367. https://doi.org/10.1145/3027385.3027450	Y	Y	Y	Y	N	N	N	D	Y	Y
28	Hui, Y. K., Mai, B., Qian, S., & Kwok, L. F. (2018). Cultivating better learning attitudes: a preliminary longitudinal study. <i>Open Learning</i> , 33(2), 155–170. https://doi.org/10.1080/02680513.2018.1454830	Y	Y	N	N	Y	N	Y	N	N	Y
29	Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. <i>Internet and Higher Education</i> , 33, 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001	Y	Y	N	N	N	N	N	N	N	Y
30	Kim, D., Park, Y., Yoon, M., & Jo, I. H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. <i>Internet and Higher Education</i> , 30, 30–43. https://doi.org/10.1016/j.iheduc.2016.03.002	Y	Y	N	N		N	Y	N	N	Y
31	Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Does time-on-task estimation matter? Implications for the validity of learning analytics findings. <i>Journal of Learning Analytics</i> , 2(3), 81–110. https://doi.org/10.18608/jla.2015.23.6	Y	Y	N	N	N	N	N	N	N	Y
32	Li, L. Y., & Tsai, C. C. (2017). Accessing online learning material: Quantitative behavior patterns and their effects on motivation and learning performance. <i>Computers and Education</i> , 114(300), 286–297. https://doi.org/10.1016/j.compedu.2017.07.007	Y	Y	N	N	N	N	N	D	N	Y
33	Lu, O. H. T., Huang, A. Y. Q., Huang, J. C. H., Lin, A. J. Q., Ogata, H., & Yang,	Y	Y	N	N	N	N	N	N	N	Y

	S. J. H. (2018). Applying learning analytics for the early prediction of students' academic performance in blended learning. <i>Educational Technology and Society</i> , 21(2), 220–232.										
34	Lukarov, V., Verbert, K., & Schroeder, U. (2019). Scaling up learning analytics in blended learning scenarios (Doctoral dissertation, Universitätsbibliothek der RWTH Aachen).	Y	Y	N	N	N	N	N	N	N	Y
35	Manzanares, M. C. S., Sánchez, R. M., Garcia Osorio, C. I., & Díez-Pastor, J. F. (2017). How do B-learning and learning patterns influence learning outcomes? <i>Frontiers in Psychology</i> , 8(MAY), 1–13. https://doi.org/10.3389/fpsyg.2017.00745	Y	Y	N	N	N	N	N	D	N	Y
36	McKenzie, W. A., Perini, E., Rohlf, V., Toukhsati, S., Conduit, R., & Sanson, G. (2013). A blended learning lecture delivery model for large and diverse undergraduate cohorts. <i>Computers and Education</i> , 64, 116–126. https://doi.org/10.1016/j.compedu.2013.01.009	M	Y	N	N	N	N	N	D	N	Y
37	Melero, J., Hernández-Leo, D., Sun, J., Santos, P., & Blat, J. (2015). How was the activity? A visualization support for a case of location-based learning design. <i>British Journal of Educational Technology</i> , 46(2), 317–329. https://doi.org/10.1111/bjet.12238	Game	Y	N	N	Y	N	N	Y	N	Y
38	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. <i>Educational Technology Research and Development</i> , 64(6), 1083–1106. https://doi.org/10.1007/s11423-016-9449-2	Video software	Y	N	N	N	N	N	N	N	Y
39	Mödritscher, F., Andergassen, M., & Neumann, G. (2013). Dependencies between E-Learning Usage Patterns and Learning Results. In <i>Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies</i> (pp. 1–8). ACM. https://doi.org/10.1145/2494188.2494206	M	N	N	N	N	N	N	N	N	Y
40	Montgomery, A. P., Mousavi, A., Carbonaro, M., Hayward, D. V., & Dunn, W. (2019). Using learning analytics to explore self-regulated learning in flipped blended learning music teacher education. <i>British Journal of Educational Technology</i> , 50(1), 114–127. https://doi.org/10.1111/bjet.12590	M	Y	N	N	N	N	N	N	N	Y
41	Musabirov, I., & Bakhitova, A. (2017). Trajectories of student interaction with learning resources in blended learning. In <i>Proceedings of the 17th Koli Calling</i>	Y	Y	N	N	N	N	N	N	N	Y

	International Conference on Computing Education Research (pp. 191–192). ACM. https://doi.org/10.1145/3141880.3141907										
42	Nespereira, C. G., Dai, K., Redondo, R. P. D., & Vilas, A. F. (2014). Is the LMS access frequency a sign of students' success in face-to-face higher education? In <i>Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '14</i> (s. 283–290). Salamanca, Spain: ACM Press. https://doi.org/10.1145/2669711.2669912	Y	Y	N	N	N	N	N	N	N	Y
43	Nguyen, Q., Rienties, B., Tempelaar, D. T., & Giesbers, B. (2016). What learning analytics-based prediction models tells us about feedback preferences of students. <i>Quarterly Review of Distance Education</i> , 17(3), 13–33.	Y	Y	N	N	N	N	N	D	N	Y
44	Nguyen, V. A., Nguyen, Q. B., & Nguyen, V. T. (2018). A Model to Forecast Learning Outcomes for Students in Blended Learning Courses Based On Learning Analytics. In <i>Proceedings of the 2nd International Conference on E-Society, E-Education and E-Technology - ICSET 2018</i> (s. 35–41). Taipei, Taiwan: ACM Press. https://doi.org/10.1145/3268808.3268827	Y	Y	N	N	N	N	N	N	N	Y
45	Nouri, J., Saqr, M & Fors, U. (2019). Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback. <i>Journal of Systemics, Cybernetics and Informatics</i> . 17(2).	Y	Y	N	N	N	N	N	D	N	Y
46	Pardo, A., Han, F., & Ellis, R. A. (2016). Exploring the relation between self-regulation, online activities, and academic performance: a case study. In <i>Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16</i> (s. 422–429). Edinburgh, United Kingdom: ACM Press. https://doi.org/10.1145/2883851.2883883	Y	Y	N	N	N	N	N	D	N	Y
47	Pardo, A., Han, F., & Ellis, R. A. (2017). Combining University Student Self-Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance. <i>IEEE Transactions on Learning Technologies</i> , 10(1), 82–92.	Y	N	N	N	N	N	N	D	N	Y
48	Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. <i>British Journal of Educational Technology</i> , 50(1), 128–138.	Y	Y	N	N	N	N	N	D	N	Y
49	Pardo, A., Mirriahi, N., Dawson, S., Zhao, Y., Zhao, A., & Gašević, D.	Video	Y	N	N	N	N	N	D	N	Y

	(2015). Identifying learning strategies associated with active use of video annotation software. In <i>Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15</i> (s. 255–259). Poughkeepsie, New York: ACM Press. https://doi.org/10.1145/2723576.2723611	software										
50	Park, Y., Yu, J. H., & Jo, I.-H. (2016). Clustering blended learning courses by online behavior data case study in a Korean higher education institute. <i>Internet and Higher Education</i> , 29, 1–11. https://doi.org/10.1016/j.iheduc.2015.11.001	Y	Y	N	N	N	N	N	N	N	N	N
51	Paskevicius, M., & Bortolin, K. (2016). Blending our practice: using online and face-to-face methods to sustain community among faculty in an extended length professional development program. <i>Innovations in Education & Teaching International</i> , 53(6), 605–615.	M	Y	N	N	N	Y	Y	D	N	N	N
52	Predić, B. (1), Rančić, D. (1), Dimić, G. (2), Štrbac, P. (2), Maček, N. (2), & Spalević, P. (3). (2018). Improving final grade prediction accuracy in blended learning environment using voting ensembles. <i>Computer Applications in Engineering Education</i> , 26(6), 2294–2306. https://doi.org/10.1002/cae.22042	Y	Y	N	N	N	N	N	N	N	N	Y
53	Snodgrass Rangel, V., Bell, E. R., Monroy, C., & Whitaker, J. R. (2015). Toward a New Approach to the Evaluation of a Digital Curriculum Using Learning Analytics. <i>Journal of Research on Technology in Education</i> , 47(2), 89–104. https://doi.org/10.1080/15391523.2015.999639	Y	Y	N	Y	Y	N	N	D	Y	Y	Y
54	Cerezo, R., Esteban, M., Sánchez-Santillán, M., & Núñez, J. C. (2017). Procrastinating behavior in computer-based learning environments to predict performance: A case study in Moodle. <i>Frontiers in Psychology</i> , 8, 1–11. doi:10.3389/fpsyg.2017.01403	Y	Y	N	N	N	N	N	N	N	N	Y
55	Rubio, F., Thomas, J. M., & Li, Q. (2018). The role of teaching presence and student participation in Spanish blended courses. <i>Computer Assisted Language Learning</i> , 31(3), 226–250. https://doi.org/10.1080/09588221.2017.1372481	Y	Y	N	Y	N	Y	N	N	Y	Y	Y
56	Saqr, M., Fors, U., & Tedre, M. (2018). How the study of online collaborative learning can guide teachers and predict students' performance in a medical course. <i>BMC Medical Education</i> , 18(1). https://doi.org/10.1186/s12909-018-1126-1	Y	Y	N	N	N	N	Y	N	N	N	Y
57	Saqr, Mohammed, Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in	Y	Y	N	N	N	N	N	N	N	N	Y

	a blended medical education course. <i>Medical Teacher</i> , 39(7), 757–767.										
58	Tempelaar, D. (1), Rienties, B. (2), & Nguyen, Q. (2). (2018). A multi-modal study into students’ timing and learning regulation: time is ticking. <i>Interactive Technology and Smart Education</i> , 15(4), 298–313. https://doi.org/10.1108/ITSE-02-2018-0015	Y	Y	N	N	N	N	N	D	N	Y
59	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. <i>Computers in Human Behavior</i> , 47, 157–167.	M	Y	N	N	N	N	N	N	N	Y
60	Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2017). Towards Actionable Learning Analytics Using Dispositions. <i>IEEE Transactions on Learning Technologies</i> , 10(1), 6–16.	M	Y	N	N	N	N	N	D	N	Y
61	Tempelaar, Dirk. (2017). How Dispositional Learning Analytics Helps Understanding the Worked-Example Principle. <i>International Association for Development of the Information Society</i> , 14.	Y	Y	N	N	N	N	N	D	N	Y
62	Tempelaar, Dirk, Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018). Student profiling in a dispositional learning analytics application using formative assessment. <i>Computers in Human Behavior</i> , 78, 408–420.	M	Y	N	N	N	N	N	D	N	Y
63	Tempelaar, Dirk, Rienties, B., & Nguyen, Q. (2018). Investigating learning strategies in a dispositional learning analytics context: the case of worked examples. In <i>Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18</i> (s. 201–205). Sydney, New South Wales, Australia: ACM Press. https://doi.org/10.1145/3170358.3170385	M	Y	N	N	N	N	N	D	N	Y
64	Tempelaar, D. (2020). Supporting the less-adaptive student: the role of learning analytics, formative assessment and blended learning. <i>Assessment & Evaluation in Higher Education</i> , 45(4), 579–593.	M	Y	N	N	N	N	N	N	N	Y
65	Van Goidsenhoven, S., Bogdanova, D., Deeva, G., Broucke, S. V., De Weerd, J., & Snoeck, M. (2020, March). Predicting student success in a blended learning environment. In <i>Proceedings of the Tenth International Conference on Learning Analytics & Knowledge</i> (pp. 17-25).	Y	Y	N	N	N	N	Y	N	N	Y
66	van Laer, S., & Elen, J. (2018). Adults’ Self-Regulatory Behaviour Profiles in Blended Learning Environments and Their Implications for Design. <i>Technology, Knowledge and Learning</i> , 1–31. https://doi.org/10.1007/s10758-017-9351-y	M	Y	N	N	N	N	N	D	N	Y

67	Van Leeuwen, A. (2019). Teachers' perceptions of the usability of learning analytics reports in a flipped university course: when and how does information become actionable knowledge? <i>Educational Technology Research and Development</i> , 67(5), 1043-1064.	M	N	N	N	N	N	N	N	N	Y
68	Villamañe, M., Álvarez, A., Larrañaga, M., Hernández-Rivas, O., & Caballero, J. (2018). Using Visualizations to Improve Assessment in Blended Learning Environments. In <i>Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM'18</i> (s. 165–169). Salamanca, Spain: ACM Press. https://doi.org/10.1145/3284179.3284209	M	Y	N	N	Y	N	N	N	N	Y
69	Whitelock-Wainwright, A., Tsai, Y. S., Lyons, K., Kaliff, S., Bryant, M., Ryan, K., & Gašević, D. (2020, March). Disciplinary differences in blended learning design: a network analytic study. In <i>Proceedings of the Tenth International Conference on Learning Analytics & Knowledge</i> (pp. 579-588)	M	Y	N	N	N	N	N	D	N	Y
70	Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. <i>Internet and Higher Education</i> , 27, 44–53. https://doi.org/10.1016/j.iheduc.2015.05.002	Y	Y	N	N	N	N	N	N	N	Y