

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success

Jelena Jovanović^{a,*}, Mohammed Saqr^{b,c}, Srećko Joksimović^d, Dragan Gašević^{e,f,g}

^a Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, 11000, Belgrade, Serbia

^b School of Computing, University of Eastern Finland, Joensuu Campus, Yliopistokatu 2, P.O. Box 111, fi-80100, Joensuu, Finland

^c EECS - School of Electrical Engineering and Computer Science, Media Technology & Interaction Design, KTH Royal Institute of Technology, Lindstedtsvägen 3, SE-100 44 Stockholm, Sweden

^d Education Futures, University of South Australia, City West Campus, 160 Currie Street, Adelaide, South Australia, 5000, Australia

^e Centre for Learning Analytics at Monash, Faculty of Information Technology, Monash University, 29 Ancora Imparo Way, Clayton, VIC, 3800, Australia

^f School of Informatics, University of Edinburgh, 10 Crichton Street, EH8 9AB, United Kingdom

^g Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia

ARTICLE INFO

Keywords:

Data science applications in education
Distance education and online learning

ABSTRACT

Predictive modelling of academic success and retention has been a key research theme in Learning Analytics. While the initial work on predictive modelling was focused on the development of general predictive models, portable across different learning settings, later studies demonstrated the drawbacks of not considering the specificities of course design and disciplinary context. This study builds on the methods and findings of related earlier studies to further explore factors predictive of learners' academic success in blended learning. In doing so, it differentiates itself by (i) relying on a larger and homogeneous course sample (15 courses, 50 course offerings in total), and (ii) considering both internal and external conditions as factors affecting the learning process. We apply mixed effect linear regression models, to examine: i) to what extent indicators of students' online learning behaviour can explain the variability in the final grades, and ii) to what extent that variability is attributable to the course and students' internal conditions, not captured by the logged data. Having examined different types of behaviour indicators (e.g., indicators of the overall activity level, those indicative of regularity of study, etc), we found little difference, if any, in their predictive power. Our results further indicate that a low proportion of variance is explained by the behaviour-based indicators, while a significant portion of variability stems from the learners' internal conditions. Hence, when variability in external conditions is largely controlled for (the same institution, discipline, and nominal pedagogical model), students' internal state is the key predictor of their course performance.

* Corresponding author.

E-mail addresses: jelena.jovanovic@fon.bg.ac.rs (J. Jovanović), mohammed.saqr@uef.fi (M. Saqr), srecko.joksimovic@unisa.edu.au (S. Joksimović), dgasevic@acm.org (D. Gašević).

<https://doi.org/10.1016/j.compedu.2021.104251>

Received 1 October 2020; Received in revised form 20 April 2021; Accepted 24 May 2021

Available online 4 June 2021

0360-1315/© 2021 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Much of the research work in Learning Analytics (LA) and Educational Data Mining (EDM) has been focused on developing predictive models of academic success and retention (Brooks & Thompson, 2017; Dawson, Joksimović, Poquet, & Siemens, 2014; Du, Yang, Shelton, Hung, & Zhang, 2019). Henceforth, the LA and EDM literature reports on a variety of analytic approaches to the creation of such predictive models, including various kinds of features and a range of statistical and machine learning algorithms that were examined in the quest for high prediction accuracy. A majority of the reported studies have been set in the context of a single course or a few courses often from the same discipline, which restricted the generalisability of the findings (Dawson, Joksimović, Poquet, & Siemens, 2019; Ifenthaler & Yau, 2020).

Early research work on predictive modeling in LA and EDM was oriented towards determining a set of ‘portable’ predictors of student performance, i.e., predictors that could be applied across different courses and institutions without the loss of predictive power (Jayaprakash, Moody, Lauría, Regan, & Baron, 2014). The leitmotif of those studies was to develop a predictive model of student success that would be independent of a particular learning context and thus could be scaled across higher education institutions (Lauría, Moody, Jayaprakash, Jonnalagadda, & Baron, 2013). Later studies pointed to and demonstrated the drawbacks of focusing on models that do not consider the specificities of the course design and disciplinary context (Conijn, Snijders, Kleingeld, & Matzat, 2017; Jovanović, Mirriahi, Gašević, Dawson, & Pardo, 2019; Rienties, Toetenel, & Bryan, 2015). For example, studies that applied predictive modelling across multiple courses, including courses from different discipline and with different instructional design (e.g., Finnegan, Morris, & Lee 2008; Gašević, Dawson, Rogers, & Gašević, 2016), tended to produce inconsistent and even conflicting findings when porting models from one course to the next (see Sect. 2.1).

The study reported in this paper aimed to further investigate the role of instructional conditions in the prediction of academic success. In particular, we wanted to empirically investigate if the conclusions from earlier studies (e.g., Finnegan et al., 2008; Gašević et al., 2016; Conijn et al., 2017) hold in a situation where courses are based on the same pedagogical underpinnings and belong to the same discipline. To that end, the current study leverages trace data from a multitude of blended medical courses, all based on problem-based learning, and examines the predictive power of several indicators of students’ engagement with the online component of the courses. Specifically, based on an extensive analysis of the related earlier studies, we define a number of indicators of online learning behaviour, including indicators of students’ activity level and regularity of study, all applicable across all the studied courses, and examine their association with the students’ learning outcomes.

In addition to examining the role of external conditions (i.e., course specific settings), we also examine the role of internal conditions, that is, factors originating from the students themselves, and to what extent such factors are accountable for the variability in the students’ course performance. This stems from the theoretical grounding of our study in the Winne and Hadwin’s model of self-regulated learning (Winne, 1996; Winne & Hadwin, 1998), which highlights the role of internal and external conditions in one’s regulation of learning and consequently, in learning achievements (see Sect. 2.2).

To attain the stated research objectives, we apply mixed effect linear regression models, to examine: i) the predictive power of indicators of students’ online learning behaviour, that is, to what extent such indicators can explain the variability in the students’ final grades, and ii) to what extent the variability in the grades is attributable to the course and students’ internal conditions, not captured by the logged data.

2. Research framework

2.1. Research background

The first wave of predictive modelling in LA and EDM was characterised by a quest for variables that could either individually or when combined allow for building accurate predictive models of student course performance or drop-out risk (Brooks & Thompson, 2017; Tempelaar, Rienties, & Giesbers, 2015). Such variables (predictors) were derived from data originating from different sources, primarily data logged by learning management systems (LMSs) and other digital systems and tools used in the learning process (e.g., Yu & Jo, 2014; Zacharis, 2015), as well as data about student demographics and/or dispositions (Shum & Crick, 2012; Tempelaar et al., 2015), and even qualitative data extracted from student essays (Strang, 2017). Most of these studies were focused on a single course or multiple offering of the same course. There have also been several studies aimed at identifying predictors common to a wide variety of formal, primarily higher education learning environments (Jayaprakash et al., 2014; Lauría et al., 2013). Such predictors were, by design, independent of contextual factors such as discipline or instructional design, so that they could be portable from one context to the next, thus allowing for a scalable use of predictive models. Common to these studies is a weak theoretical grounding, that is, a lack of reliance on learning science theories for predictor selection and/or explanation of the resulting models (Gašević, Dawson, & Siemens, 2015; Gašević et al., 2016; Lust, Collazo, Elen, & Clarebout, 2012).

More recently, LA research has started to draw on learning theories and recognise the relevance of theory based predictive modelling (Gašević et al., 2016; Rienties et al., 2015). In particular, a theoretically grounded approach to predictive modelling, and learning analytics in general, i) allows for leveraging existing knowledge in the learning sciences; ii) enables meaningful research design and interpretation of the results, and iii) provides grounds for re-examining existing and building new knowledge (Gašević et al., 2015; Rogers, Gašević, & Dawson, 2015; Wise & Shaffer, 2015).

Contemporary learning theories consider learning as intrinsically situated (Lave & Wenger, 1991). For example, constructivist theories emphasize the interplay of instructional context and student internal conditions (Winne, 1996; Winne & Hadwin, 1998), whereas process theories accentuate the dialectic between instruction and learning (Engeström, 2014). Sociocultural approaches

recognise the role of social context and individual differences in the ways this context is internalised (Nolen & Ward, 2008). By grounding their work in contemporary learning theories, several recent studies demonstrated the weaknesses of general (context-agnostic) predictive models and the relevance of considering learning settings (Conijn et al., 2017; Finnegan, Morris, & Lee, 2009; Gašević et al., 2016; Joksimović, Gašević, Loughin, Kovanović, & Hatala, 2015; Jovanović, Mirriahi, et al., 2019; Kizilcec, Reich, Yeomans, Dann, Brunskill, Lopez et al., 2020). For instance, Finnegan et al. (2009) examined the association between several indicators of student online learning behaviour and academic achievement across 22 courses in 3 broad academic disciplines (English and Communication; Social Sciences; and Math, Science, and Technology). They found a notable diversity in indicators' predictive power across different courses. Furthermore, no indicator proved to be a significant predictor of academic achievement across the investigated disciplines. In a later study that included nine courses, some of which differed in discipline and instructional design, Gašević et al. (2016) found that the predictive power of the same behavioural indicators varied even among courses from the same disciplinary area. This study was replicated by Conijn et al. (2017), using a larger sample (17 courses) of more homogenous courses, both in terms of the subject matter and instructional design (hence, also the LMS usage pattern). They found no comprehensive set of variables that could consistently predict student performance across multiple courses.

The importance of instructional context for predictive modelling has also been recognized in studies that explored the association of students' online interactions and regularity of study, on one hand, and their course performance, on the other. For example, Joksimović et al. (2015) used learning trace data from a total of 204 offerings of 29 online graduate courses to investigate different interaction types as defined in a contemporary theory of distance and online education. Their findings confirmed the relevance of the course context for predicting students' academic achievement, and indicated that it may be even more relevant than students' individual differences. Jovanović et al. (2019a) compared predictive models with generic (i.e., course-design-agnostic) vs. course-design-specific indicators, using trace data from three consecutive offerings of a blended course with a flipped classroom design. The predictive models with generic indicators were able to explain only a small portion of the overall variability in the students' course performance, and were significantly outperformed by models with course specific indicators.

To sum up, the literature on predictive modelling in LA and EDM indicates that models aimed at predicting students' academic achievements tend to be context dependent, which prevents their portability from one learning context to another. In particular, the predictive power of such models proved to be heavily influenced by the course discipline and the enacted instructional design. What is less known, and thus is the focus of the current study, is whether such models can retain predictive power when ported across courses with highly similar contexts - namely courses in the same discipline and with the same or highly similar instructional design.

2.2. Theoretical grounding of the study

The current study is theoretically grounded in the constructivist, metacognitive approach to self-regulated learning (SRL) proposed by Winne and Hadwin (Winne, 1996; Winne & Hadwin, 1998). This model posits that learners are active agents who employ a range of cognitive, physical, and digital tools to process raw information in order to create learning artefacts and progress towards their learning objectives. To that end, learners regulate their learning processes by continuously evaluating the quality of their learning products (against defined standards) and the effectiveness of the chosen study tools and tactics. This process of metacognitive monitoring is influenced by a range of internal and external conditions. The former includes, for example, learners' motivation, prior knowledge, and affective states. External conditions are largely determined by elements of the instructional settings such as the teacher's role, course requirements, and availability and form of feedback.

This SRL model is relevant to our study given our focus on blended learning and the well-established relevance of self-regulation of learning in online and blended settings (Broadbent & Poon, 2015; Rasheed, Kamsin, & Abdullah, 2020). Furthermore, SRL skills are of particular relevance when external conditions are determined by a loosely structured instructional design where learners are given much autonomy in how they approach learning tasks, as is the case in our study where all courses are based on problem-based learning (English & Kitsantas, 2013).

In the present study, we consider both internal and external conditions that affect the learning process. Regarding external conditions, we focus on their major component, namely instructional settings when defining variables (indicators) for the predictive models as well as when interpreting the results of such models. The role of instructional conditions in learners' self-regulation efforts and their use of learning tools has been examined and evidenced in numerous studies (e.g., Garrison & Cleveland-Innes, 2005; Gašević, Mirriahi, Dawson, & Joksimović, 2017; Lust et al., 2012). Accordingly, we can expect that the instructional settings would shape the students' interaction with the online component of a blended course, i.e., their use of the learning resources (i.e., materials, tools, activities) made available through the institutional LMS (Rienties et al., 2015). Based on this, we can hypothesize that students will interact more frequently and intensively with resources that are either directly or indirectly recommended by the instructional design (e.g., use a discussion forum to collaborate on a shared task). Earlier studies (e.g., Jovanović, Dawson, Joksimović, & Siemens, 2020; Jovanović, Mirriahi, et al., 2019; Panzarasa, Kujawski, Hammond, & Roberts, 2016) have demonstrated a positive association between the students' learning outcomes and their level and regularity of interaction with the resources relevant to the given instructional conditions. Considered from the perspective of predictive modelling, this implies that variables indicative of students' level and regularity of interaction with instructionally relevant course elements can be expected to be highly predictive of the students' academic achievement.

The effect of internal conditions on the overall learning process and learning outcomes has been demonstrated in several studies (e.g., Kintu, Zhu, & Kagambe, 2017; Manwaring, Larsen, Graham, Henrie, & Halverson, 2017). In this study, due to the restrictions imposed by data availability, we do not examine individual elements of students' internal conditions, but consider their impact in aggregate. That is, after controlling for a variety of external conditions, we examine the effect of students, as individuals, on the

learning outcomes, and in particular, the extent to which individual differences cause variability in the learning outcomes. In their comprehensive review of Winne and Hadwin's SRL model, [Greene and Azevedo \(2007\)](#) highlight the importance of students' perceptions and subjective experience of the learning context for the overall learning process, including the moderating role that such perceptions have on other internal conditions. Considering students' individual differences, which may stem from their prior knowledge of the topic under study, metacognitive knowledge, and motivation ([Winne, 1996](#)), it is reasonable to expect that different students would act differently in the same learning settings. Thus, we can hypothesize that such individual differences, by impacting students' learning behaviour, can constitute an important source of variability in the students' learning outcomes. Differences caused by subjective experience of the shared learning space has also been emphasized by sociocultural approaches ([Nolen & Ward, 2008](#)) that "quite explicitly argue for the salience of the individual experience in a shared environment" ([Martin et al., 2015](#), p. 27). Accordingly, a substantial variance can be expected between students as different events impact them, since students perceive and process those events in idiosyncratic ways.

To sum up, our theoretical framework postulates the importance of internal and external conditions for regulation of learning and consequently the learning outcomes. These propositions have been proved by prior empirical studies. The current study aims to examine the predictive power of internal and external conditions across a set of homogenous courses, and the extent to which the variability in the students' learning outcomes can be explained by each group of conditions.

2.3. Research questions

This study builds on the methods and findings of related earlier studies to further explore factors predictive of students' academic success in blended learning settings. In doing so, it differentiates itself by (i) relying on a larger and more homogeneous course sample (15 courses, 50 course offerings in total), and (ii) considering both internal and external conditions as factors affecting the learning process. As for the former, we controlled for several sources of heterogeneity among courses; namely, our study sample consists of courses that originate from the same institution, belong to the same discipline (medicine), and are based on the same nominal learning design (blended problem-based learning). In addition, all courses are of small to medium size (50 students, on average), which further contributes to the reduction in variability and differentiates our course sample from those used in previous studies (mostly large courses with several hundred students).

Based on all the above, the study is guided by the following two research questions:

- RQ1: *What trace-based behaviour indicators (variables) and to what extent can explain the variability in the academic success, measured through the final course grade, in multiple blended courses that belong to the same discipline and share the same nominal instructional approach?*
- RQ2: *To what extent external conditions (namely instructional context) can explain the variability in the final course grades and to what extent that variability is explained by internal conditions (i.e., an interplay of individual student features)?*

3. Methods

3.1. Context

The data for the study originate from several blended courses taught at the same university between 2014/15 and 2017/18 academic years. The courses are a part of an undergraduate medical curriculum that integrates different basic medical topics under themed courses - for example, the course on growth and development includes physiology, histology, pathology, and biochemistry of growth and development, as well as clinical applications. The online part of the curriculum is technically supported by a Moodle learning management system (LMS).

The curriculum relies on problem-based learning (PBL) as the general instructional approach, which is applied in a blended learning mode. This blended PBL process can be summarised as follows: the learning objectives of each course form the basis for the weekly problems. The work on a weekly problem starts on the first day of the week with a face-to-face session where students, split into groups (7–10 students), are offered an open-ended scenario (i.e., *problem*) to serve as a stimulus for actively seeking information about what needs to be learnt to solve the given *problem*. Having set their learning objectives for the week, the students continue interacting online, via discussion forums. Each group has a separate discussion forum, facilitated by the tutor assigned to the group. During the week, group members discuss the problem and share resources, explanations, and solutions. The last day of the week is scheduled for the second face-to-face session where the groups discuss their solutions and conclusions.

The final grade is the sum of PBL and exam grades. The PBL grades make up 20% of the final grade and are assessed by a tutor according to the student's contribution to the PBL weekly discussions. The rest of the final grades come from the written exams (multiple choice and short essay questions) which tests knowledge acquisition of the PBL objectives that were the subject of the online interactions.

3.2. Data source

The main source of data for the study is the log database of the Moodle LMS that served as the common online platform for all the courses in the medical curriculum. As a part of data preprocessing, the original event types, as captured in Moodle logs, were mapped into less granular, learning-related actions. This was done since the original event types were too granular and low-level (e.g., resource-

	Label	Description	Studies that used the same / similar indicators*
Overall Level of Activity (OLA)	OLA_1	Session count, normalised	Gašević et al., 2016; Gitinabard et al., 2019; Zacharis, 2015
	OLA_2	Total session length (in seconds), normalised	Gitinabard et al., 2019; Tempelaar et al., 2015; Zacharis, 2015
	OLA_3	Average (median) session length (in seconds), normalised	Conijn et al., 2017; Gitinabard et al., 2019
	OLA_4	Average (median) number of learning actions per session, normalised	Gitinabard et al., 2019; Jovanović et al., 2020
	OLA_5	Proportion of active days	Jovanović et al., 2020
	OLA_6	Average (median) number of active days per week	Jovanović et al., 2020
	OLA_7	Average (median) time distance between two consecutive active days	
	OLA_8	Proportion of active weeks	Saqr et al., 2017
Learning Action specific Level of Activity (LALA)	LALA_1 - LALA_4	Proportion of active days per learning action type Four indicators in total, one for each learning action type: <i>course_main_viewed</i> , <i>lecture_viewed</i> , <i>forum_consume</i> , and <i>forum_contribute</i>	
	LALA_5 - LALA_8	Proportion of active weeks per learning action type Similar to the previous set: four indicators, one for each type of learning action	
Overall Regularity of Study (ORS)	ORS_1	Entropy of learning action counts per session	
	ORS_2	Entropy of session length	Conijn et al., 2017
	ORS_3	Proportion of weeks when the student was active on the first day of the week	Saqr et al., 2019
	ORS_4	Average (median) proportion of learning actions on the first day of the week	
Learning Action specific Regularity of Study (LARS)	LARS_1 - LARS_4	Entropy of daily action counts per learning action type (four indicators, one for each type of learning action)	Jovanović et al., 2019a
	LARS_5 - LARS_8	Entropy of weekly active day counts per learning action type (four indicators, one for each type of learning action)	
*Some indicators (e.g., session count or total time online) were used in a large number of studies; in such cases, only a sample of studies is listed			

Fig. 1. Overview of the indicators used in the study.

view, forum-search) for a meaningful analysis. Therefore, we defined a mapping from the ‘original’, low-level logged events to a small set of derived, higher-level (less granular) learning actions:

- forum contribution (*forum_contribute*) - an action of contributing to a discussion forum by either posting, updating or deleting a message, or uploading a resource to share
- forum consumption (*forum_consume*) - an action of browsing or reading forum content
- grade-report view - an action of viewing/examining one’s own grades or overall course grades
- access to learning materials (*lecture_viewed*) - an action of accessing and viewing individual learning resources or directories with learning resources
- access to the main course page (*course_main_viewed*) - an action of accessing the course landing page with relevant course information, links, and updates.

In addition to the aforementioned, some other (low-level) event types were logged in the Moodle database. However, these were either unrelated to the student interaction with the LMS (teacher related events) or were very infrequent and thus were not suitable for the analysis.

After the preprocessing, data for 15 courses and 50 offerings in total (each course with 3–4 offerings) were kept for the study, based on the following selection criteria: at least 3000 learning actions of any type and at least 500 forum consumption actions throughout the course duration. These criteria were set to assure a minimum level of LMS use in a course and a minimal adherence to the PBL design, respectively.

The second source of data was the student information system that was used to gather course grades data. Since courses differed in the grades value range, grades were standardised at the course level, using median and interquartile range, since in some of the courses grades were not normally distributed and had a few outliers. Students who did not take the final exam were excluded from the analyses.

The final sample included 2472 enrolments, and 803,179 learning actions. As the students were enrolled in multiple courses, the sample included 344 unique students, who were enrolled, on average, in 6 courses from the final sample. Table A (in the Supplementary file) provides some basic descriptive statistics about the number of students, active course weeks, and logged actions per course offering in the final sample. The table gives statistics for the overall action counts as well as the counts per different action types (as defined above).

To simplify the narration, from here on, we use the term *course* to refer to individual course offering and use the terms *course* and *course offering* when the distinction between the two is important.

3.3. Indicators of learning behaviour

A set of indicators of the students’ online learning behaviour were derived from the (preprocessed) Moodle log data. The indicators were defined by examining variables used in related studies and the available data, i.e., logged events in the study sample. Note that in examining the literature our focus was on indicators used in blended and online higher education contexts, excluding massive open online courses (MOOCs), considering numerous differences between MOOC and formal education settings, especially when examined from the perspective of student success prediction (Joksimović, Poquet, Kovanović, Dowell, Mills, Gašević et al., 2017). In addition, we restricted our focus on indicators that could be derived from data logged by LMSs. That is, we did not consider variables that could be extracted from data collected by custom built learning tools or extensions of LMSs that allow for custom data collection. This decision was motivated not only by our study context, but primarily by the fact that a large majority of fully online and blended courses use LMSs as the primary and often only learning platform (Bouchrika, 2020).

To assure a comprehensive coverage of various aspects of students’ online learning behaviour, we defined four categories of indicators (Fig. 1): i) general indicators of the overall level of students’ interaction with the online portion of the course; ii) indicators of the activity level in relation to different online learning resources available through the LMS; iii) general regularity of study indicators; iv) indicators of regularity of engagement with different kinds of online learning resources. Since the courses in the study sample differed notably in terms of their duration (see n_{weeks} in Table A), to neutralize this difference, all indicators were normalised, that is, instead of absolute value indicators, relative value (e.g., proportions) indicators were defined. In general, the use of relative indicators increases the generalisability of predictive models (Hung, Shelton, Yang, & Du, 2019).

Several indicators are related to learning sessions. A session is identified as a continuous sequence of learning actions where the time gap between any two consecutive actions is below the chosen threshold. The threshold for each course is set to the 85th percentile of the time gaps between two successive learning actions within the given course. To account for differences in session length across courses, session-related activity level indicators (*OLA_1–4*) were normalised by dividing their value with the maximum value for the indicator in the given course.

An active day is defined as a day with at least one learning action of any type. Proportion of active days indicator (*OLA_5*) is computed by dividing the number of active days by the length of the course expressed in days. This and the other two indicators based on active days (*OLA_6–7*) were introduced as means of capturing, from different ‘angles’, the extent of students’ ‘presence’ in the online component of the course. We chose this type of indicators, over often used event counts (e.g., Conijn et al., 2017; Zacharis, 2015), to neutralize the effect of quick succession of clicks (i.e. sequences of almost instantaneous events) and students’ tendency to exhibit non-uniform, bursty temporal patterns in interaction with online learning resources (Sagr, Fors, & Tedre, 2017).

An active week is defined as a week with the number of active days equal to or above the average number of active days per week in the course. Proportion of active weeks indicator (*OLA_8*) is the ratio of the number of active weeks and the course length expressed in

weeks. The reason for having this and other weekly based indicators (*LALA_5–8*, *LARS_5–8*), in addition to the daily based ones, is that all courses had weekly periodicity: in each week, there was a new problem (i.e. learning task) to tackle (see Sect. 3.1) and it was important for students to follow problem-related announcements and materials in each week as they formed the basis for all learning activities.

Indicators *LALA_1–8* are defined for different types of learning actions derived from LMS logs (Sect. 3.2). An active day for a learning action type is a day with at least one action of the given type (e.g., an active forum-contribution day is a day when a student had at least one forum-contribution action). A week is considered active for a particular type of learning action, if the student had at least two active days related to that type of learning action. These features were introduced as they allow for examining the level of students' interaction with the course activities at a lower level of granularity than the one offered by overall activity indicators (*OLA_5–7*), while still neutralising non-uniform, bursty temporal patterns often present in students' interaction with online learning resources.

Similar to the split of the activity level indicators into more general ones (*OLA_1–8*) and those related to specific types of learning actions (*LALA_1–8*), we have also considered two groups of regularity of study indicators, defined at different levels of granularity: less granular ones reflecting overall regularity of online learning behaviour (*ORS_1–4*) and more granular ones indicating regularity in relation to particular types of learning actions (*LARS_1–8*).

Regularity indicators based on the first day of week (*ORS_3–4*) were motivated by the work of Saqr, Nouri, and Fors (2019) who found that early in the week participation was a consistent predictor of high achievement.

Several regularity of study indicators (*ORS_1*, *ORS_2*, *LARS_1–8*) are based on entropy, as a measure of the uniformity of a discrete probability function. For example, in the case of the *ORS_1* indicator, probabilities were estimated by dividing the number of learning actions in each session with the total number of learning actions (across all sessions) for the given student in the given course. Thus obtained probabilities were used in the formula for Shannon's entropy (Shannon, 1948). Entropy reaches its maximum value when all the probabilities are uniformly distributed, that is, in the context of *ORS_1*, when the counts of learning actions per session are the same. This means that higher entropy implies higher regularity of study and vice versa, making entropy suitable for estimating regularity of learning-related behaviour (Panzarasa et al., 2016).

3.4. Data analysis: mixed effect linear models

Mixed-effect models combine fixed and random effects and thus allow for assessing the association between the fixed effects and the dependent variable after accounting for the variability due to the random effects (Hayes, 2006). These models have been used in several similar studies (e.g., Gašević et al., 2017; Joksimović et al., 2015; Jovanović, Gašević, Pardo, Dawson, & Whitelock-Wainwright, 2019) to estimate the strength of association between some constructs of interest (modelled as fixed effect(s) and the outcome) while controlling for the variability that originates in the individual differences among students and/or courses (random effects). Accordingly, we used mixed-effect linear models to examine the association between indicators of students' engagement with online learning activities (fixed effects) and their course performance (outcome), while at the same time accounting for the variability due to the student and course specific features (random effects). In doing so, we were able to address our two research questions (RQs): models' performance and coefficients associated with fixed effects allowed for addressing RQ1, whereas the proportion of variability associated with the random effects was used to address RQ2.

Mixed effect linear models were built with the final course grade as the dependent variable. In all models, two random effects were used - student (*user_id*) and course offering (*course_id*) - modelled as partially crossed random effects. They were partially crossed as not all students took all the courses (on average, each student attended 6 courses).

Indicators of the students' engagement with the online portion of the courses (Fig. 1) were used as fixed effects. In particular, we have created four models (M_{OLA} , M_{LALA} , M_{ORS} , M_{LARS}), one using each group of indicators defined in Fig. 1 as fixed effects. Having identified significant predictors in each model, we used those predictors, as fixed effects, to build the final model (M_{FINAL}). This approach was motivated by our objective to examine the cross-course predictive power of different types of learning behaviour indicators and in each group identify those that were the most stable (i.e. consistent) across courses.

Each full model, i.e., model with fixed-effects included, was compared to the null model, i.e., model with random effects only, to determine if the fixed effects had significant association with the students' course grades above and beyond the random effects. The comparison was based on Akaike Information Criterion (AIC), Log Likelihood (LL), and a likelihood ratio test.

To further evaluate the models, but also to answer RQ2, we have computed, for each model, pseudo R^2 and Intraclass Correlation Coefficient (ICC). Pseudo R^2 method (Nakagawa & Schielzeth, 2012) is used to estimate the proportion of variance in the dependent variable explained by a mixed effect model. In particular, in case of mixed-effects models, two kinds of R^2 statistics are relevant: i) marginal R^2 (R_M^2) that quantifies the variance explained by the fixed effects, and ii) conditional R^2 (R_C^2) that represents the variance explained by the entire model (i.e., both random and fixed effects). ICC is used to estimate the proportion of the total variance in the outcome that was to be attributed to each of the random effects, namely student and course.

It should be also mentioned that the assumptions for linear mixed effect models - including linearity, normality and homoscedasticity of residuals, and absence of multicollinearity and influential points - were verified for all the models. To address the problem with linearity, the outcome variable (i.e., course final grades) was transformed using nonparanormal transformation (Liu, Laerty, & Wasserman, 2009), implemented in the *huge* R package (Jiang, Fei, Liu, Roeder, Lafferty, Wasserman et al., 2020). To avoid the problem of multicollinearity, from each model, we removed indicators with Variance Inflation Ratio (VIF) above 3 (Zuur, Ieno, & Elphick, 2010). We have also checked for influential points using the Cook's distance (Cook, 1977). Having inspected the identified

potentially influential observations, we found that they offset each other's influence and thus did not affect the models.

The significance level of 0.05 was used in all statistical tests. All the computations were done in the R programming language, using *lme4* (Bates, Mächler, Bolker, & Walker, 2015), *merTools* (Knowles, Frederick, & Whitworth, 2020), and *performance* (Lüdtke, Makowski, Waggoner, & Patil, 2020) R packages.

4. Results and discussion

Results of mixed effect models built for each group of indicators (Fig. 1) are summarised in Table 1. The results for the final model, the one built using only significant predictors from each indicator group, are given in Table 2. Only indicators that proved significant are reported. We first examine results related to our first research question (RQ1), by presenting and discussing the models' performance and their fixed effects (i.e., trace-based behaviour indicators) that proved to be significant predictors of the final course grades. This is followed by an examination of the random effects, to address our second research question (RQ2).

4.1. Research question 1

The comparisons of the full models (i.e., models with both fixed and random effects included) with the corresponding null models (i.e., models with random effects only) demonstrated that the full models yielded a significantly better fit (Table B in the Supplementary file). To make the narration more compact, in the following we refer to full models as models.

All models had similar predictive power, with conditional R^2 (R_C^2) ranging from 0.717 to 0.726, that is, the overall model explaining around 72% of the variability in the students' final course grade. This suggests that the examined groups of indicators (Fig. 1) are comparable in their ability to capture elements of students' online behaviour that tend to affect learning outcomes. Furthermore, the examined indicators tend to be mutually 'competitive', instead of complementing one another. This is evident in the high correlations among the indicators, and the absence of additive effect of predictors in the final model.

The portion of variability explained by fixed effects (R_M^2) was low in all models, ranging from 0.026 (in M_{ORS}) to 0.047 (in M_{FINAL}). In other words, only 3–5% of variability in the students' final grades could be explained by the indicators of learning activity level or regularity of study. This may be attributed to the fact that we were not able to capture students' interaction that happened outside the institution's LMS, either through other online communication channels or in the face-to-face portion of the courses. These results can also be, at least partially, explained by the fact that even though the courses from the sample were based on the same pedagogical model (PBL), there was a difference in the enactment of this model in different courses. This is evident in the extent of use of different online learning resources across the courses (Table A in the Supplementary file). Accordingly, the relevance of behaviour-based indicators and the predictive power of models based on them could have been expected to differ across the courses from the sample, as confirmed by our closer investigation (given in Section C of the Supplementary file). This further suggests that the use of the same pedagogical model is not a guarantee of the portability of predictive models across courses; differences in the enactment of a pedagogical model bring in the heterogeneity across courses that eventually impairs attempts at having portable predictive models. In other words, the complex interplay of various factors (e.g., instructors' and students' characteristics) tends to lead to variability in the actual

Table 1

Results for the mixed effect models with the four groups of indicators defined in Fig. 1 as fixed effects; only significant indicators are shown.

Model	Predictors	Coeff.	St. Error	St. coeff.	p-value	
M_{OLA}	OLA_1 - Session count, normalised	0.419	0.133	3.152	0.0017	
	$R_C^2 = 0.717$	OLA_2 - Total session length, normalised	0.380	0.104	3.641	0.0003
	$R_M^2 = 0.046$	OLA_4 - Average number of learning actions per session, normalised	0.542	0.245	2.218	0.0273
	$ICC_S = 0.682$	OLA_8 - Proportion of active weeks	0.279	0.069	4.050	<0.0001
	$ICC_C = 0.022$					
M_{LALA}	LALA_1 - Proportion of active days related to the <i>lecture_viewed</i> learning actions	0.709	0.141	5.034	<0.0001	
	$R_C^2 = 0.726$	LALA_2 - Proportion of active days related to the <i>forum_consume</i> learning actions	0.592	0.179	3.312	0.001
	$R_M^2 = 0.030$					
	$ICC_S = 0.689$					
	$ICC_C = 0.029$					
M_{ORS}	ORS_2 - Entropy of session length	0.134	0.022	6.145	<0.0001	
	$R_C^2 = 0.725$	ORS_3 - Proportion of weeks when the student was active on the first day of the week	0.212	0.062	3.420	0.0006
	$R_M^2 = 0.026$					
	$ICC_S = 0.692$					
	$ICC_C = 0.026$					
M_{LARS}	LARS_1 - Entropy of daily counts of the <i>lecture_viewed</i> learning actions	0.057	0.026	2.188	0.0287	
	$R_C^2 = 0.725$	LARS_3 - Entropy of daily counts of the <i>forum_contribute</i> learning actions	0.101	0.029	3.520	0.0005
	$R_M^2 = 0.033$	LARS_5 - Entropy of weekly counts of the <i>lecture_viewed</i> active days	0.072	0.028	2.581	0.01
	$ICC_S = 0.683$	LARS_8 - Entropy of weekly counts of the <i>course_main_viewed</i> active days	-0.131	0.028	-4.724	<0.0001
	$ICC_C = 0.032$					

Table 2
Results for the final mixed effect linear model; only significant indicators are shown.

	Predictors	Coeff.	St. Error	St. coeff.	p-value
M_{FINAL}	OLA_2 - Total session length, normalised	0.374	0.092	4.077	0.0005
$R_C^2 = 0.717$	LARS_3 - Entropy of daily counts of the <i>forum_contribute</i> learning actions	0.068	0.029	2.322	0.0205
$R_M^2 = 0.047$	LARS_5 - Entropy of weekly counts of the <i>lecture_viewed</i> active days	0.070	0.025	2.777	0.0055
$ICC_S = 0.675$	LARS_8 - Entropy of weekly counts of the <i>course_main_viewed</i> active days	-0.074	0.027	-2.703	0.0070
$ICC_C = 0.028$					

application of a pedagogical model, thus negatively affecting the replicability of prediction results. This result provides an analytic confirmation of the conclusion made by Dawson and colleagues, based on a review of the last ten years of LA research, that “an explanation for the lack of replicable outcomes lies in the dynamics of the systems within which LA operates.” (Dawson et al., 2019, p.8).

In the model with indicators of the overall activity level as fixed effects (M_{OLA} , Table 1), four indicators proved to have statistically significant positive association with the outcome. Those are normalised session count (*OLA_1*), normalised total session length (*OLA_2*), average number of learning actions per session (*OLA_4*), and proportion of active weeks (*OLA_8*), that is, proportion of weeks with average or above average number of active days. Positive coefficients of all four predictors confirm that higher level of activity is positively associated with the course success.

The model with learning action specific activity indicators (M_{LALA} , Table 1) showed that only two out of eight indicators had significant effect on the final course grades: (i) proportion of days when a student accessed learning materials at least once, i.e., had at least one action of the *lecture_viewed* type (*LALA_1*); and (ii) proportion of days with at least one access to the discussion forum for the sake of browsing or reading forum content (*LALA_2*). Both indicators were positively associated with the outcome, and the former (*LALA_1*) proved to be one of the strongest predictors across the four models (based on standardised coefficients, Table 1).

In the model with overall regularity of study indicators (M_{ORS} , Table 1), two indicators proved to be significant predictors of the outcome variable. With its standardised coefficient of 6.154, entropy of session length (*ORS_2*) turned out to be the strongest predictor across the four models, suggesting that having learning sessions of consistent length is associated with higher course performance. Another significant indicator (*ORS_3*) suggests that those students who are active on the first day of week tend to have better course grades.

The model with learning action specific regularity of study indicators (M_{LARS} , Table 1) identified four indicators as having significant effect on the final course grade. The positive coefficient associated with the entropy of daily counts of *forum_contribute* actions (*LARS_3*) indicates that higher regularity in posting to discussion forums tended to be associated with higher final course grades. Likewise, for the consistency in the lecture materials access, be it the consistency of the daily counts (*LARS_1*) or the consistency in the counts of active days per week (*LARS_5*), there was a positive association with the final grades. On the other hand, regularity in the weekly counts of days with at least one access to the course landing page (*course_main_viewed*) (*LARS_8*) proved to have a negative association with the grades.

The final model (M_{FINAL} , Table 2) gathered significant predictors from the above described four models. However, a few predictors (*OLA_1*, *ORS_1*, *LARS_4*) had to be removed to avoid multicollinearity. Among the remaining predictors, four proved significant:

- The overall time spent online (*OLA_2*), a general indicator of the overall level of activity in relation to the online portion of the course; it has been identified as a relevant predictor in several previous studies (e.g., Gitinabard, Xu, Heckman, Barnes, & Lynch, 2019; Tempelaar et al., 2015; Zacharis, 2015).
- Regularity in the daily counts of discussion forum postings (*LARS_3*); its relevance could be attributed to the PBL design of the sample courses where students' communication and resource sharing through discussion forums had a key role in attaining weekly learning objectives; related studies have primarily examined the frequency of forum postings (e.g., Conijn et al., 2017; Saqr et al., 2017), far less considered regularity of posting.
- The significance of the weekly regularity indicators, in particular, those reflecting regular access to the course landing page (*LARS_8*) and the lecture materials (*LARS_5*), can be explained by the weekly periodicity of the curriculum, which made it important for students to follow announcements and materials related to weekly problems as they formed the basis of lectures, practicals and exams. The negative association between regular access to the course main page and the student learning performance can be explained by the presence of students who were regular in accessing the course, but not in much else (i.e., did not sufficiently engage with the learning resources offered through the LMS). Indicators of regular weekly use of online learning resources were also proven significant in a recent study that examined the predictive power of regularity of pre-class activities in a similar learning context (flipped classroom with weekly schedule) (Jovanović, Mirriahi, et al., 2019). Being based on a significantly larger number of courses and course editions, the current study offers a stronger confirmation of the results reported by Jovanovic and colleagues (2019a).

4.2. Research question 2

To address RQ2, we examined the random effects of the mixed effect models, and in particular, the value of the Intraclass Correlation Coefficient (ICC) of each random effect in each model. ICC measures the proportion of the total variance in the outcome variable that can be attributed to a particular random effect. Since we had two crossed random effects, student and course offering, we computed ICC for student-level grouping (ICC_s) and ICC for the grouping at the course offering level (ICC_c). The results reported in Tables 1 and 2 show that across all the models, the variability accounted by the two random factors, after controlling for the student behaviour in the course (i.e., fixed effects), was rather consistent. In particular, student characteristics, not captured by the indicators of their online learning behaviour, accounted for about 68% of variability in the final course grades, whereas the course offering specificities explain only about 3% of the variability. ICC can also be interpreted as the correlation among observations within the same group. This interpretation suggests that there is a high correlation among observations, across different courses, related to the same student. On the other hand, there is a very low correlation among observations corresponding to different students within the same course offering. These results suggest that a student's internal conditions explain a large proportion of the variability in their learning outcomes. This is consistent with the theoretical underpinnings of our study as well as the findings of previous related studies.

From the theoretical perspective, according to the Winne and Hadwin's (1998) model of SRL (see Sect. 2.2), decisions that learners make about their learning are influenced by internal and external conditions. By considering students' interaction with the LMS (fixed factors in our models) and the course specificity not reflected in logs (course offering random effect), we are capturing, to an extent, the impact of instructional design, that is, external conditions. An aggregate of internal conditions is in our model captured by the student random effect, which proved to carry the largest explanatory power and thus requires further investigation.

The important role of internal conditions that needs to be further explored was also observed by Conijn et al. (2017). Having identified that a high proportion of variance in students' final grades could be explained at the student level, Conijn and colleagues noted that "none of the usage characteristics that have been used in the literature before seemed to pick up this variance" (Conijn et al., 2017, p.27). Similar findings were reported by Gašević et al. (2016) who examined student personal characteristics across several courses, and found a notable difference among the students. However, those characteristics that were predominantly demographic (e.g., age, gender, proxies for socio-economic status) had less explanatory power than trace-based predictors, when used in models predicting students' course performance. This implied that individual level features (i.e., internal conditions) that contributed to the variability in the course performance were not the demographic ones. The novel insight that our study brings to the literature is that even when several sources of heterogeneity among courses are controlled for (same institution, same discipline, and same nominal learning design), differences in learning outcomes are only marginally explained by students' online learning behaviour, whereas the primary source of variability is in their internal conditions.

5. Limitations

This study was by its design focused on a set of homogenous courses – all from the same institution, all in the same discipline, and all based on the same pedagogical model. While this homogeneity was a necessity for adequately addressing our research questions, it is also an impediment to the generalisability of the study results. Hence, our findings need to be verified in other contexts, including different nominal pedagogical approaches, different cultures, and different disciplines.

Another limitation stems from the trace data source being restricted to the institution's Moodle log database, which is rather restrictive in the contextual data of the logged events. In particular, for each learning-related event, relevant contextual data include user id, timestamp, event type, and URL of the related course page. This, consequently, imposed restrictions on the kinds of indicators we were able to use in predictive models. So, even though we have defined and examined a broad range of behaviour indicators (Fig. 1), many of which were used in previous studies, we were not able to examine, for example, those that would further reflect the specificities of group interaction such as various network-based indicators (Saqr, Fors, & Nouri, 2018). Thus, it would be relevant to examine if the study findings would be different if additional indicators, especially those more closely reflecting the underlying pedagogical model, were used.

Finally, the examined courses were based on the blended learning mode with important face to face sessions at the beginning (familiarizing with the problem and setting learning objectives) and end (discussing problem solutions) of each course week. Since we only had data about the online component of the course, we were not able to account for the students' offline behaviour and thus might have missed some important predictors. Still, this is not a limitation specific to our study but common to all related studies.

6. Implications

Our findings indicate the need to focus on individual, student-level analytics, in addition to the current focus towards analytics of the overall cohort. This is in line with Winne's proposal (2017) for leveraging big data and analytics for a new approach to research in learning science. Specifically, Winne suggests that traces collected at the individual level ($N = 1$ or $N = me$) can be used for analysing, understanding, and advancing learning processes of individual students. Similarly, Wilson, Watson, Thompson, Drew, and Doyle (2017) pointed out the diversity of learners' approaches to study and online spaces, and the need for considering the uniqueness of learning processes of individual learners. For example, traces of individual students can be used to study learning tactics and strategies (Matcha, Gašević, Uzir, Jovanović, & Pardo, 2019), time management (Ahmad Uzir et al., 2019), stability of SRL (Winne & Perry, 2000) and other aspects of self-regulatory learning of individuals.

The study results also provide robust empirical evidence for the need to "embed" learning design in learning traces, that is, to

instrument online learning systems and tools with means for capturing instructionally relevant data. This would allow for measuring constructs of relevance for learning and teaching instead of or in addition to the currently omnipresent system-administration oriented logs. This would further empower LA to provide data, methods, and tools for continuous iterative validation and improvement of learning design (Lockyer, Heathcote, & Dawson, 2013). The relevance of capturing pedagogically meaningful trace data for developing learning analytics that are both scalable and contextualised (i.e., aligned with the learning design) has also been argued by Shibani, Knight, and Shum (2019). In particular, their Contextualizable Learning Analytics Design (CLAD) model includes features that should be “pedagogically salient, and not simply technical representations of low-level trace data” (Shibani et al., 2019, p.212), which may require an extension of the underlying technical infrastructure for data gathering and processing. The value of pedagogically rich learning logs has been demonstrated, for example, by Jovanović, Mirriahi, Gašević, Dawson, and Pardo (2019) who made use of custom-made learning logs, informed by the instructional design of the given course, to create course-specific indicators that proved much more predictive of students’ course performance than general, course-design-agnostic indicators. Probably, the best example of how such instrumentation of a learning environment can be done and the benefits it may bring about have been demonstrated by Winne et al. (2019) with their nStudy learning tool that allows for gathering data about students’ cognitive, metacognitive, and motivational processes as students work on their study tasks.

Furthermore, there is a need to collect data about students’ internal factors. Some earlier studies relied on surveys and similar self-reporting tools to collect such data (e.g., Kintu et al., 2017; Lee, Park, & Davis, 2018). However, considering the advantages of trace-based measures over traditional self-reporting instruments (Winne & Jamieson-Noel, 2002; Zhou & Winne, 2012), a better approach could be to instrument learning environments with tools that would allow for seamless eliciting of students’ perceptions, opinions, and affective states during the learning process. For example, Jovanović, Gašević, Pardo, Dawson, and Whitelock-Wainwright (2019) were able to seamlessly collect students’ perceived cognitive load and self-efficacy throughout the learning process using a self-reporting tool integrated in the LMS. An alternative or a complementary approach may be to employ the experience sampling method (Zirkel, Garcia, & Murphy, 2015) to quickly and seamlessly elicit students’ opinions, perceptions, affective states, or mood. For example, Manwaring et al. (2017) used experience sampling to collect data about students’ cognitive and emotional engagement throughout a semester and examine factors that affected the level of engagement. Martin et al. (2015) examined longitudinal data about students’ motivation and engagement, and after controlling for socio-demographics factors and prior achievement found substantial within-day (intraindividual) and between-student variability in motivation and engagement.

Considering the increasing use of black-box prediction models (e.g., various forms of deep learning models) and the ongoing interest in developing portable predictive models of academic success and retention (e.g., Gitinabard et al., 2019; Hung et al., 2019), it is relevant to consider the difference between purely predictive and explanatory models. The distinction between the two modeling cultures is well explained by Bergner (2017) and Shmueli (2010) and could be summarised as follows: in explanatory modelling, the focus is on obtaining the most accurate representation of the underlying theory, whereas predictive modelling seeks to maximise empirical precision, occasionally at the expense of theoretical accuracy. The current study and majority of the cited related studies (Sect. 2.1) can be qualified as explanatory. An example of the predictive modeling approach is a recent study by Hung et al. (2019) who have demonstrated that a combination of two stage prediction models and the use of relative behaviour variables (general, frequency-based activity level indicators) can lead to generalisable predictive models of at-risk students. However, Hung et al. (2019) have provided very limited information about the pedagogical underpinnings of the sampled courses. Furthermore, their study was neither grounded in learning theory nor it explained the results from a theoretical perspective. Though models of this type can be relevant, if we want to provide informed assistance to students (e.g., by offering personalised feedback) and/or teachers (e.g., by enabling validation of the deployed instructional design), we need to focus on explanatory models and the insights such models can offer. This was well articulated by Rosé, McLaughlin, Liu, and Koedinger (2019) who argued for an approach to learning analytics that provides accurate predictions, alongside actionable insights that contribute to both learning science and educational practice.

7. Conclusions

The reported study examined a comprehensive set of features covering several aspects of students’ online learning behaviour, including indicators of activity level and regularity, at diverse levels of granularity (e.g., course as a whole or individual learning activities). The study was based on a sample of 15 blended courses (50 course editions) that were homogeneous in terms of the institutional settings, discipline, nominal learning design, and course size. The study contributions can be summarised as follows:

- The same nominal pedagogical model and study setting do not guarantee the portability of predictive models across courses; heterogeneity caused by differences in the enactment of the pedagogical model negatively affects the models’ cross-course predictive power.
- The overall time spent online, regular daily contributions to the discussion forum, and regular weekly access to the course and lecture materials proved to be the most significant predictors of students’ learning outcomes (measured through the final course grade) across courses. While the relevance of the former predictor has already been well established in the literature, the latter two have been far less studied and their cross-course significance is a novelty that the current study brings to the literature.
- The examined indicators of online behaviour are not complementary, that is, when used together they do not explain significantly more variability in students’ learning outcomes than when each indicator group is used individually for predictive modelling. In other words, there is an upper limit in the variability that can be explained by manifested learning behaviour.
- Students’ internal conditions explain a large proportion of the variability in their learning outcomes.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compedu.2021.104251>.

Credit author statement

Jelena Jovanović: Conceptualization, Methodology, Formal analysis, Investigation, Validation, Writing – all stages, Visualisation, Mohammed Saqr: Conceptualization, Methodology, Investigation, Data curation, Writing – review & editing, Srećko Joksimović: Conceptualization, Methodology, Writing – review & editing, Dragan Gašević: Conceptualization, Methodology, Writing – review & editing, Supervision

References

- Ahmad Uzir, N., Gašević, D., Matcha, W., Jovanović, J., Pardo, A., Lim, L.-A., et al. (2019). Discovering time management strategies in learning processes using process mining techniques. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Transforming learning with meaningful technologies* (pp. 555–569). Springer International Publishing. https://doi.org/10.1007/978-3-030-29736-7_41.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bergner, Y. (2017). Measurement and its uses in learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *The handbook of learning analytics* (pp. 34–48). Alberta, Canada: Society for Learning Analytics Research (SoLAR). Retrieved from <http://solaresearch.org/hla-17/hla17-chapter1>.
- Bouchrika, I. (2020, June 30). *50 online education statistics: 2020 Data on higher learning & corporate training*. Guide2research. <https://www.guide2research.com/research/online-education-statistics>.
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Brooks, C., & Thompson, C. (2017). Predictive modelling in teaching and learning. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *The handbook of learning analytics* (pp. 61–68). Alberta, Canada: Society for Learning Analytics Research (SoLAR).
- Conijn, R., 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>
- Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15–18.
- Dawson, S., Gašević, D., Siemens, G., & Joksimović, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 231–240). New York, NY, USA: ACM. <https://doi.org/10.1145/2567574.2567585>.
- Dawson, S., Joksimović, S., Poquet, S., & Siemens, G. (2019). Increasing the impact of learning analytics. In *Proc. Of the international conference on learning analytics and knowledge*. <https://doi.org/10.1145/3303772.3303784>. Tempe, Arizona, USA, March 2019 (LAK'19).
- Du, X., Yang, J., Shelton, B. E., Hung, J.-L., & Zhang, M. (2019). A systematic meta-Review and analysis of learning analytics research. *Behaviour & Information Technology*, 1–14. <https://doi.org/10.1080/0144929X.2019.1669712>, 0(0).
- Engeström, Y. (2014). *Learning by expanding: An activity-theoretical approach to developmental research*. Cambridge, UK: Cambridge University Press.
- English, M. C., & Kitsantas, A. (2013). Supporting student self-regulated learning in problem- and project-based learning. *Interdisciplinary Journal of Problem-Based Learning*, 7(2). <https://doi.org/10.7771/1541-5015.1339>. Available at:
- Finnegan, C., Morris, L. V., & Lee, K. (2009). Differences by course discipline on student behavior, persistence, and achievement in online courses of undergraduate general education. In , *Vol. 10. Journal of college student retention: Research, theory and practice* (pp. 39–54). <https://doi.org/10.2190/CS.10.1.d.1>.
- Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education*, 19(3), 133–148. https://doi.org/10.1207/s15389286ajde1903_2
- Gašević, D., Dawson, S., Rogers, T., & Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, 28, 68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>
- Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64–71.
- Gašević, D., Mirriahi, N., Dawson, S., & Joksimović, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior*, 67, 207–220. <https://doi.org/10.1016/j.chb.2016.10.026>
- Gitinabard, N., Xu, Y., Heckman, S., Barnes, T., & Lynch, C. F. (2019). How widely can prediction models be generalized? Performance prediction in blended courses. *IEEE Transactions on Learning Technologies*, 12(2), 184–197. <https://doi.org/10.1109/TLT.2019.2911832>
- Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research*, 77(3), 334–372. <https://doi.org/10.3102/003465430303953>
- Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research*, 32(4), 385–410. <https://doi.org/10.1111/j.1468-2958.2006.00281.x>
- Hung, J.-L., Shelton, B. E., Yang, J., & Du, X. (2019). Improving predictive modeling for at-risk student identification: A multistage approach. *IEEE Transactions on Learning Technologies*, 12(2), 148–157. <https://doi.org/10.1109/TLT.2019.2911072>
- Ienthaler, D., & Yau, J. Y. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research & Development*, 68, 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>
- Jayaprakash, S. M., Moody, E. W., Lauria, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47.
- Jiang, H., Fei, X., Liu, H., Roeder, K., Lafferty, J., Wasserman, L., et al. (2020). *huge: High-Dimensional undirected graph estimation (1.3.4.1)*. <https://CRAN.R-project.org/package=huge>.
- Joksimović, S., Gašević, D., Loughin, T. M., Kovanović, V., & Hatala, M. (2015). Learning at distance: Effects of interaction traces on academic achievement. *Computers & Education*, 87, 204–217. <https://doi.org/10.1016/j.compedu.2015.07.002>
- Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., et al. (2017). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*. <https://doi.org/10.3102/0034654317740335>
- Jovanović, J., Dawson, S., Joksimović, S., & Siemens, G. (2020). Supporting actionable intelligence: Reframing the analysis of observed study strategies. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 161–170. <https://doi.org/10.1145/3375462.3375474>
- Jovanović, J., Gašević, D., Pardo, A., Dawson, S., & Whitelock-Wainwright, A. (2019b). Introducing meaning to clicks: Towards traced-measures of self-efficacy and cognitive load. *Proc. of the 9th International Conference on Learning Analytics & Knowledge*, 511–520. <https://doi.org/10.1145/3303772.3303782>
- Jovanović, J., Mirriahi, N., Gašević, D., Dawson, S., & Pardo, A. (2019a). Predictive power of regularity of pre-class activities in a flipped classroom. *Computers & Education*, 134, 156–168. <https://doi.org/10.1016/j.compedu.2019.02.011>
- Kintu, M. J., Zhu, C., & Kagambe, E. (2017). Blended learning effectiveness: The relationship between student characteristics, design features and outcomes. *International Journal of Educational Technology in Higher Education*, 14(1), 7. <https://doi.org/10.1186/s41239-017-0043-4>

- Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., et al. (2020). Scaling up behavioral science interventions in online education. In *Proceedings of the national academy of sciences*. <https://doi.org/10.1073/pnas.1921417117>
- Knowles, J. E., Frederick, C., & Whitworth, A. (2020). *Tools for analyzing mixed effect regression models*. CRAN. <https://cran.r-project.org/web/packages/merTools/index.html>.
- Lauria, E. J. M., Moody, E. W., Jayaprakash, S. M., Jonnalagadda, N., & Baron, J. D. (2013). Open academic analytics initiative: Initial research findings. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 150–154. <https://doi.org/10.1145/2460296.2460325>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge university Press.
- Lee, J., Park, T., & Davis, R. O. (2018). What affects learner engagement in flipped learning and what predicts its outcomes? *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.12717>. n/a(n/a).
- Liu, H., Laerty, J. D., & Wasserman, L. (2009). The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *Journal of Machine Learning Research*, 10, 2295–2328.
- Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, 57(10), 1439–1459.
- Lüdecke, D., Makowski, D., Waggoner, P., & Patil, I. (2020). *performance: Assessment of regression models performance*. CRAN. <https://doi.org/10.5281/zenodo.3952174>. R package <https://easystats.github.io/performance>.
- Lust, G., Collazo, N. A. J., Elen, J., & Clarebout, S. (2012). Content management systems: Enriched learning opportunities for all? *Computers in Human Behavior*, 28(3), 795808.
- Manwaring, K. C., Larsen, R., Graham, C. R., Henrie, C. R., & Halverson, L. R. (2017). Investigating student engagement in blended learning settings using experience sampling and structural equation modeling. *The Internet and Higher Education*, 35, 21–33. <https://doi.org/10.1016/j.iheduc.2017.06.002>
- Martin, A. J., Papworth, B., Ginns, P., Malmberg, L.-E., Collie, R. J., & Calvo, R. A. (2015). Real-time motivation and engagement during a month at school: Every moment of every day for every student matters. *Learning and Individual Differences*, 38, 26–35. <https://doi.org/10.1016/j.lindif.2015.01.014>
- Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., & Pardo, A. (2019). Analytics of learning strategies: Associations with academic performance and feedback. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 461–470. <https://doi.org/10.1145/3303772.3303787>
- Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142.
- Nolen, S. B., & Ward, C. J. (2008). Sociocultural and situative approaches to studying motivation. In M. L. Maehr, S. A. Karabenick, & T. C. Urdan (Eds.), *Advances in motivation and achievement*. 15 (pp. 425–460). Bingley, UK: JAI Press.
- Panzarasa, P., Kujawski, B., Hammond, E. J., & Roberts, C. M. (2016). Temporal patterns and dynamics of e-learning usage in medical education. *Educational Technology Research & Development*, 64(1), 13–35. <https://doi.org/10.1007/s11423-015-9407-4>
- Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A systematic review. *Computers & Education*, 144, 103701. <https://doi.org/10.1016/j.compedu.2019.103701>
- Rienties, B., Toetenel, L., & Bryan, A. (2015). “Scaling up” learning design: Impact of learning design activities on LMS behavior and performance. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, 315–319. <https://doi.org/10.1145/2723576.2723600>
- Rogers, T., Gašević, D., & Dawson, S. (2015). Learning analytics and the imperative for theory driven research. In C. Haythornthwaite, R. Andrews, J. Fransma, & E. Meyers (Eds.), *The SAGE handbook of E-learning research* (2nd ed.). London, UK: SAGE Publications Ltd.
- Rosé, C. P., McLaughlin, E. A., Liu, R., & Koedinger, K. R. (2019). Explanatory learner models: Why machine learning (alone) is not the answer. *British Journal of Educational Technology*, 50, 2943–2958. <https://doi.org/10.1111/bjet.12858>
- Saqr, M., Fors, U., & Nouri, J. (2018). Using social network analysis to understand online Problem-Based Learning and predict performance. *PLoS One*, 13(9), Article e0203590. <https://doi.org/10.1371/journal.pone.0203590>
- 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>
- Saqr, M., Nouri, J., & Fors, U. (2019). Time to focus on the temporal dimension of learning: A learning analytics study of the temporal patterns of students' interactions and self-regulation. *International Journal of Technology Enhanced Learning*, 11(4), 398–412.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shibani, A., Knight, S., & Shum, S. B. (2019). Contextualizable learning analytics design: A generic model and writing analytics evaluations. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 210–219. <https://doi.org/10.1145/3303772.3303785>
- Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310. <https://doi.org/10.1214/10-STS330>
- Shum, S. B., & Crick, R. D. (2012). Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 92–101. <https://doi.org/10.1145/2330601.2330629>
- Strang, K. D. (2017). Beyond engagement analytics: Which online mixed-data factors predict student learning outcomes? *Education and Information Technologies*, 22(3), 917–937. <https://doi.org/10.1007/s10639-016-9464-2>
- Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, 47, 157–167.
- Wilson, A., Watson, C., Thompson, T. L., Drew, V., & Doyle, S. (2017). Learning analytics: Challenges and limitations. *Teaching in Higher Education*, 22(8), 991–1007. <https://doi.org/10.1080/13562517.2017.1332026>
- Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, 8(4), 327–353. [https://doi.org/10.1016/S1041-6080\(96\)90022-9](https://doi.org/10.1016/S1041-6080(96)90022-9)
- Winne, P. H. (2017). Leveraging big data to help each learner and accelerate learning science. *Teachers College Record*, 119(3), 1–24.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, 27(4), 551–572. Oct. 2002.
- Winne, P. H., & Perry, N. E. (2000). Chapter 16—measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50045-7>
- Winne, P. H., Teng, K., Chang, D., Lin, M. P.-C., Marzouk, Z., Nesbit, J. C., et al. (2019). nStudy: Software for learning analytics about processes for self-regulated learning. *Journal of Learning Analytics*, 6(2). <https://doi.org/10.18608/jla.2019.62.7>, 95–106–195–106.
- Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, 2(2), 5–13. <https://doi.org/10.18608/jla.2015.22.2>
- Yu, T., & Jo, I.-H. (2014). Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 269–270. <https://doi.org/10.1145/2567574.2567594>
- Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, 27, 44–53. <https://doi.org/10.1016/j.iheduc.2015.05.002>
- Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, 22(6), 413–419. <https://doi.org/10.1016/j.learninstruc.2012.03.004>
- Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher*, 44(1), 7–16. <https://doi.org/10.3102/0013189X14566879>
- Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3–14. <https://doi.org/10.1111/j.2041-210X.2009.00001.x>