LMS log activity as a predictor of learning success on an undergraduate flipped classroom course of cellular biology

Vesa Paajanen[1]

[1] *Department of Environmental and Biological Studies, University of Eastern Finland, Finland*

Abstract

Monitoring student activity in learning management system (LMS) can provide useful information for learning design, student support, and development of LMS analytical tools. This study used LMS log information to detect the elements which help students to reach their learning goals in a flipped classroom course of bioscience, identify students' time management difficulties, and develop tools to detect students with risk of drop out during the course. A traditional lecture course of bioscience was modulated into flipped classroom, and Moodle activity of 100 voluntary student was analyzed in the study to reveal the potential pitfalls of the course structure. Effectiveness of course content to increase learning success was tested with pair-vise comparisons, variance analyzes and best subset regression. Although in general, the students were able to reach the learning goals better after the modulation, LMS log information revealed several challenges in self-regulated learning. The learning activity of the students with high drop-out/failure risk was highly periodical, for which 25 % of their LMS use was happening in a single day of 7-week course. Beside the time management challenges, use of elements of resources and supportive mechanisms affected course passing and grading statistically significantly. These parameters were together able to predict the learning success (R^2 = 0,645) much better than student's earlier studies or main study subject. Interestingly, detail analyses of single elements revealed small but significant, nonlinear effects of self-evaluation and external www-links for low and high-grade students, respectively. LMS use of students at-risk was significantly different already during the first week of the course, which indicate high possibilities of learning analytics to identify students who need help during the course.

Keywords

learning analytics, higher education, Log activity, flipped classroom, self-regulated learning

1. Introduction

During the last decades, higher education has been turning online to make the learning available for students outside the campus. A dramatic example of this shift happened during the pandemic 2020, during which 1,6 billion students were forced to learn outside their classrooms. This emergency online-teaching period introduced all teachers to online facilities, but also increased isolation and unsatisfaction of students as well as teacher's anxiety for the teaching in online platforms. At the same time, analytical information of student retention and completion rates as well as microscale learning analytics has become an important target for university and national scale development of higher education [1]. Thus, there is a large need for information about student supportive online learning environments, and factors that could be used to identify and guide students at risk of drop-out [2].

Recently, lack of social interaction has been commonly used as an argument against the online learning. Teachers feel unable to track the learning processes when lecturing online, or when the students are making their studies independently in learning management systems (LMS) like Moodle.

Proceedings of the Finnish Learning Analytics and Artificial Intelligence in Education Conference (FLAIEC22), Sept 29–30, 2022, Joensuu, Finland
EMAIL: vesa.paajanen@uef.fi

The benefits of social interaction and teacher's role to chop the learning process into small, reachable parts has been known for almost 100 years [3]. However, noninteractive lecturing has been widely used also in pre-covid times, especially in the STEM [4], which can explain why students prefer online participation also when the lecture rooms have been re-opened.

Teaching practices to support students have more than one variable: studies are not only online or onsite but also synchronous or asynchronous, interactive or independent, contain formative or summative assessments etc. Moreover, blended teaching (e.g. flipped classroom), in which students' personal (online) and group (lecture room) time alternate, has become widely used method to reach students outside the campus and to increase social interactions in teaching [5]. However, as teachers can have unsatisfactory experiences from online teaching, there is need for studies in which efficiency of selected methods to support learning has been demonstrated and/or elements of course design linked to learning success and/or risk of drop-out has been identified. The latter goal require learning analytics and it can provide useful information for faculty members modifying courses, identifying, and supporting students at-risk, and developing LMS analytical tools.

2. Theoretical background

During the last 20 years, learning design has been widely used to identify the good teaching practices [6]. As each course is different and may contain dozens of actions, learning design use categorization of course elements into for example resources, tasks, and supportive mechanisms. The importance of elements or categories to help students to reach the learning outcomes can be tested with several methods. Beside the self-report studies, newer research has focused more on students' act on LMS with learning analytics [7]. Learning analytics does not only provide numeric information of students' actions in LMS but also new opportunities to monitor students' learning progress [8]. When students are operating with any online platform, their activity is tracked for mostly security and/or commercial reasons. Their activity on LMS is saved as log data which can be used to estimate the factors of learning success [9,10]. This provides educators information on the usefulness of course materials, students at-risk, and time of study work. However, analyses of LMS Log activity require mining of large data sets, beside which it can have ethical challenges [11].

Since the beginning of learning analytics era, it has become clear that some, but not all, elements predict student retention and academic performance [12]. Benefits of several individual parameters and categories on learning success has been demonstrated. For example, use of resources (e.g. video lectures, texts) sound obvious requirement for reaching the learning outcomes and there is correlation between LMS use and academic performance [13,14]. However, parameters like time spent on video lectures have controversial effect on student performance [10,15]. Tasks increase student activity and their benefits for learning improvement has been well demonstrated especially in case of quizzes [16–19]. Recently, supportive mechanisms have been widely studied as interaction-based learning like Community of Inquiry CoI [20] has become more popular in online education. Interactivity prevents isolation in online learning [21,22] and can increase student satisfaction and learning outcomes as well as reduce the drop-out rate [23–25]. Supportive mechanisms can also lead to better learning outcomes [24,26,27].

As online and blended courses can assume study work outside the lecture hours, students must be able to plan their learning schedules and follow their plans. This can cause time management challenges. When studying independently, students may use unoptimized techniques [28] beside which their skills in self-regulated learning can be affected by the social background [29]. Therefore, students need training for independent study work, and differences in their learning schedules can have strong influence on learning success. For example, periodical activity of course material in LMS demonstrate challenges with time management [30–33] which will lead to procrastination and last night LMS activity [34–36]. This can be revealed only with data mining of Moodle log information.

Success on the course depends on the combination of elements (e.g. resources, tasks, supportive mechanisms) in which students are participating during their study work and time of this activity. Moreover, LMS can provide information on the differences in course material use between withdrawers and completers with different learning success, and differences in learning strategies among students [7,10,13]. The LMS log data has been used to estimate best-practice models for online learning [8,37] and the purpose for learning analytics of LMS log data is to identify and help students-at-risk as early as possible [2].

Beside students' activity on the course, their demographic and static data can determinate learning success [2]. During the pandemic 2020, there was more concern at first-year students than students with "academic skills". Therefore, students with academic background could have less difficulties than freshmen to pass their studies. Likewise, students with similar studies in their previous education could reach the learning outcomes easier than students who take the course as a part of their minor studies.

2.1 Flipped classroom

Flipped Classroom (FL) has become a common method to engage students with interactivity without limiting the study work in a classroom. It is based on delivery of information which students will study independently before participating interactive (usually onsite) elements [38]. This individual and group space combining method can help interactivity on large-scale and multi-campus courses and its efficacy on learning outcomes has been demonstrated [for example 39]. Students participating FL course will be more responsible for their learning for which they require more self-regulation skills but co-work in interactive elements will set short-term goals for their study work [40]. Moreover, as students are familiar with the course material, they are able to reach more challenging tasks (e.g. problem-based learning, experiment-design) together with their peers during the group space, and ask applied questions from their teacher in the meetings.

Tracking the activity of each student before they will participate on the interactive elements is challenging in large FC courses. However, working with peers can increase students' motivation to study the material in advance. In many cases, majority of learning in FC courses happens in LMS for which the learning can be organized also fully online, if the interactive elements are voluntary or organized with online meetings. This can be essential for students living far away from the campus, but at the same time it can lead into lower learning outcomes for students studying without these supportive mechanisms.

As several elements of FL and online are similar, there is no differences in the studying strategies of these courses [41]. Therefore, LMS can provide information of student actions and best practices related to learning success in FL courses. Moreover, when the supportive mechanisms are voluntary and course can be completed fully online, importance of interactivity to learning success can be tested.

In the current study, I have used learning analytics to detect the activity of students on a new flipped classroom course. The course was *modulated* from traditional lecture course into self-regulated studying behavior demanding flipped classroom teaching on spring 2016 as a part of flipped classroom project in the university [42,43]. During the modulation 90 min lectures were chopped into 10 to 30 min segments and recorded as screencasts. Lectures were delivered in Moodle environment together with separate quiz segments, external links, and guidance for interactive elements (group work & chat).

3. Aims of the study

The study was using LMS Log information to detect the elements that help students to reach their learning goals in flipped classroom course of bioscience and student actions related to high risk of drop-out and course failures. The study was focusing on the following questions:

1) What happens on students' learning success when lecture course is changed into FL?
2) Are students able to manage their study time in distant, self-regulated learning?
3) How different resources, tasks, and supportive mechanisms help students to reach the learning outcomes of the course?
4) Is it possible to identify students-at-risk and how early the identification based on LMS log information can be made?

4. Methods

In fall semester 2016 145 students from University of Eastern Finland and 16 students of continuous learning were participating on the Basics of Cellular and Molecular Biology course. Majority (57 %) of the students were freshmen, for who the course was one of their first experience on Higher Education, beside which 50 % of the students were taking the course outside their major studies. Therefore, students had large variance on background knowledge and limited information, how to study successfully on university level courses.

The course contained textbook (Campbell Biology 2011 pp 78-475), short video lectures, podcast series (in Soundcloud) and external www-links as source of information, quizzes for self-evaluation of learning and weekly 2 h group work sessions and online-chat (Slack) to build the learning community and to synchronize learning during the 8 Cp and 7-week course. The course material was available in Moodle LMS from the beginning of each study week after which it was in a free use for the students. Students were encouraged to use quizzes and to participate on live group work and online chat with extra points (<10 % of the maximal points) for their activity. The extra points were based on the participation – not on the skills of the students on quizzes or group work outcomes.

The course started 24th October and ended 9th December 2016. Written examinations were held in three segments (Biochemistry 11/7, Cell biology 11/28 & Molecular biology 12/16) in Fall 2016 to evaluate the learning outcomes of the students. In Spring 2017 students were able to participate on two examinations during which they were able to remake any of the segments. Grading was based on 0-5 scaling in which 50 % of the maximal points (150 p) were required for passing the course with a grade 1/5 and 90% were required for achieving the highest grade 5/5 on the course. The grading was based on similar evaluation criteria as was used in the prior traditional lecture course to demonstrate the learning of students in blended/online learning. Questions in the examinations were based on the course material but students had free choice on the material they wanted to use during the course. Therefore, students were encouraged to blended participation and self-regulated learning before participating on group work, but they were allowed to study completely online and without using any components of the course.

Table 1
Potential course outcome affecting factors

Category	Parameter
Static data	Main study subject Years of study in the university Sex
Resources	Amount of all LMS logs including all activity in Moodle Visiting events in pages containing any studying material of the course Visiting events in pages containing lecture videos Visiting events in pages containing external www-links
Tasks	Visiting events in pages containing quizzes
Time management	Days of any LMS activity Average LMS log per day with any activity Days in which LMS activity is higher than the average activity of the student
Supportive mechanisms	Participation and activity of the student on online-chat Participation and activity of the student on group work sessions

Academic research in which students' LMS activity is combined with their learning outcome have legal and ethical challenges. Therefore, a separate agreement was asked from the students to participate the study, in which their anonymized LMS activity and learning outcomes were analyzed to reveal the behavioral differences affecting on learning success. Majority of the students (100 persons) gave permission for the use of their information on this study. The course outcome and participation on course activities were combined with analyzes of Moodle log information of the voluntary students. Several literature-based parameters [2,6] were selected as potential course outcome affecting factors, beside student static data (Table 1). As over 90% of the students were

Caucasian with Finnish nationality and Finnish as their mother language, demographic information was not used in this study.

4.1. Analyses

Log data containing 62900 events were analyzed with SigmaPlot 14.0 (Systat Software) in three setups. Effects of flipped classroom on studying behavior and course outcome were analyzed by comparing the students with participation on social activity (online-chat, group work session) with students studying the course completely online and without need for interactivity. These two group comparisons were made either with Student's t-test or Mann-Whitney rank sum test depending on the normality of the distribution. To evaluate the best practices of the studying behavior, potential factors of course outcome were analyzed for students with outcome (grades 0 – 5) and these groups were compared by using Kruskal-Wallis One Way Analysis of Variance on Rank together with Dunn's Method as a post-hog method to calculate the difference between the individual groups. To detect the main components of the factors affecting course outcome regression between course points (0-150) and the factors were analyzed with Best Subset Regression in which linear regression of parameter combinations will be used to model the course outcome. In all cases, $P<0{,}05$ was considered as statistically significant difference between the groups and in Best Subset regression the simplest model in which least bias (Cp is among the smallest), adjusted R^2 highest and the variance inflation factor (VIF) below 5 in all parameters, was selected [44].

5. Results

In general, shift from 46 h lecture course to a flipped classroom mode with 20 h of video lectures and 14 h of group work activities was successful: failure rate decreased from 20,4 to 14,4 % at the same time when proportion of excellent grades (5/5) increased from 5,4 to 15,2 % of the students. Moreover, course examinations were planned to be at least as demanding as on the prior lecture course on Bloom's taxonomy. There were no statistically significant differences in learning success between students selecting the course as their major studies, taking the course in their minor studies, or participating on the course as continuous learners. In comparison, freshmen were able to achieve the learning goals better than students with earlier university studies ($P=0{,}003$). Therefore, lecture course experience in university studies does not help students in more self-oriented study work and flipped classroom can be used even for the newcomers in the higher education without reduction of course learning goals and outcomes.

Electronic material and LMS provide more information on the studying behavior of the students than is possible to collect on a traditional lecture course. The LMS use was clearly periodical (Fig 1) during the course and students were producing over 960 events during the most active day but only 116 events on the average. Typically, LMS activity was high on the day before each examination and similar peak was seen also in the podcast channel related to the course (Fig 1A).

Periodical activity on LMS indicate a link between the student outcome and the last night study work. Therefore, the average daily LMS activity of the students with different grades was tested (Fig 1B). Students either failing the course or passing it below 60 % learning outcome (grade 1/5) were studying mainly on a single day just before each examination (days 16, 36 & 55). The main difference between the students falling or passing with a low grade was on the daily activity: students passing the course with low grade were using the LMS twice the average activity of the students failing the course. In comparison, the students with better outcome on the course (Grades 2 – 5) were studying on several days and in the highest groups with above 80 % learning outcome (Grades 4 & 5) the peak before exams was absent.

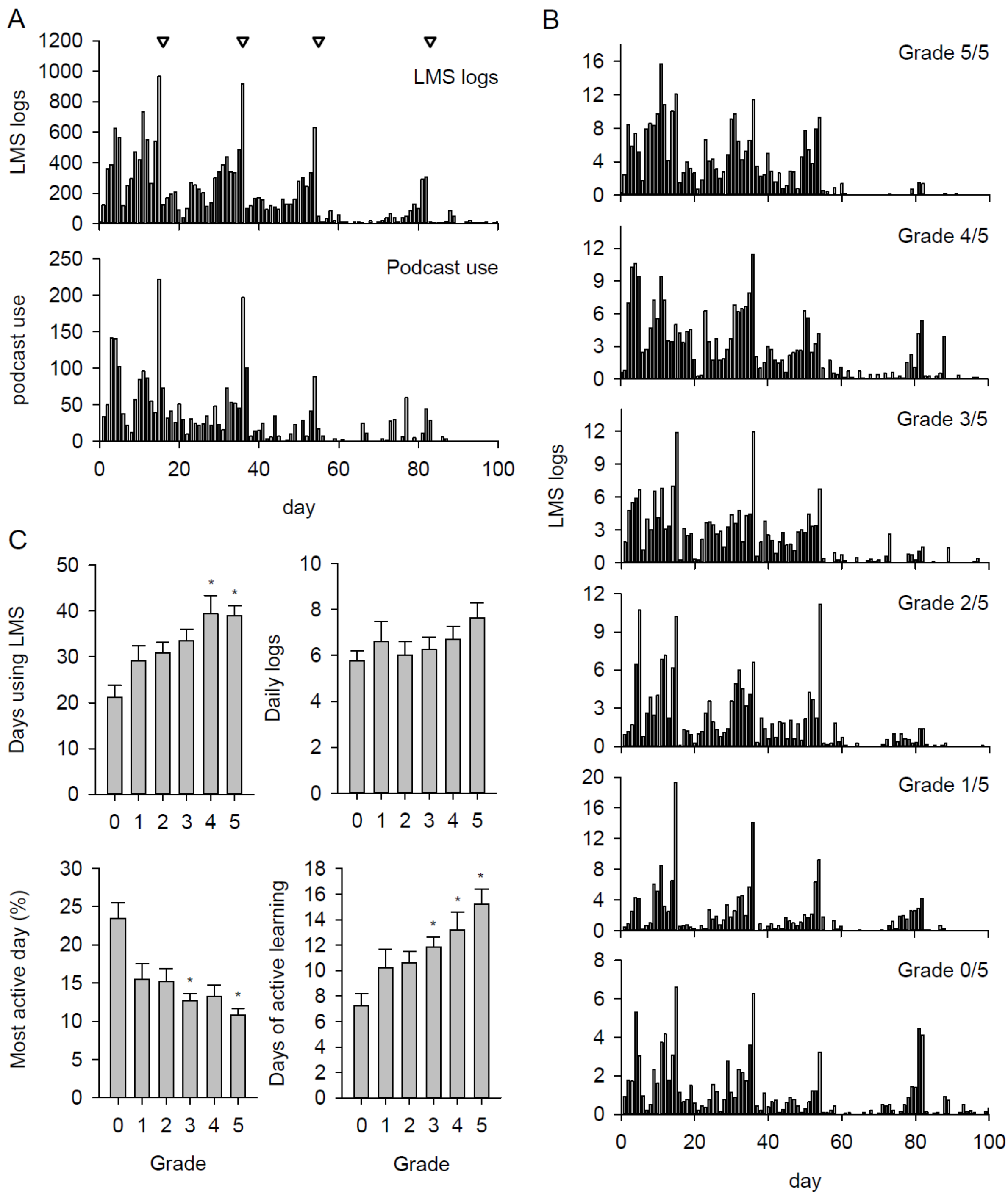

Figure 1: Periodic activity of the LMS. A) All logs on the LMS during the first 100 days of the course (upper figure). All activity of students not participating on the study are excluded. Days of examinations are marked with white triangles on the top. Listening events on the course podcast channel was available also outside the students of this study (bottom). B) Use of any learning components in LMS made by students with different course outcome. C) Days of any activity (Logs) on LMS in different student groups (up left), average daily LMS use (up right), proportion of LMS use of the most active day of all activity during the course (bottom left) and number of days with more LMS activity than the average of the student's daily use (bottom right). Asterix (*) indicate statistically significant difference between students passing the course with certain grade and the students failing the course (0).

The regularity and periodical activity on the course were tested quantitatively with four parameters (Fig 1C). Firstly, the number of days with any activity on the LMS was higher in students with high learning outcomes (upper left). Secondly, the number of log events on each day indicating the amount of daily study hours was similar in all student groups (upper right, P=0,243). Thirdly, the last night studying was detected by comparing the activity of the most active day to all logs of the student on LMS (lower left). Students failing the course were using LMS only on single days for which their logs on the most active day was almost ¼ of all activity on the course whereas students passing the course were studying more regularly. Fourth parameter mark the regularity of study work with the number of days during which students have been using LMS more than their personal average use. This parameter does not count days of minimal activity (few minutes logging on LMS) like the days of any LMS activity for which it can detect better the number of days in which students have actively used LMS for their learning. All these statistical comparisons indicate that time management is a critical element of studying success. Students failing the course are using less days for their active learning. However, the amount of daily LMS use is not a key to success.

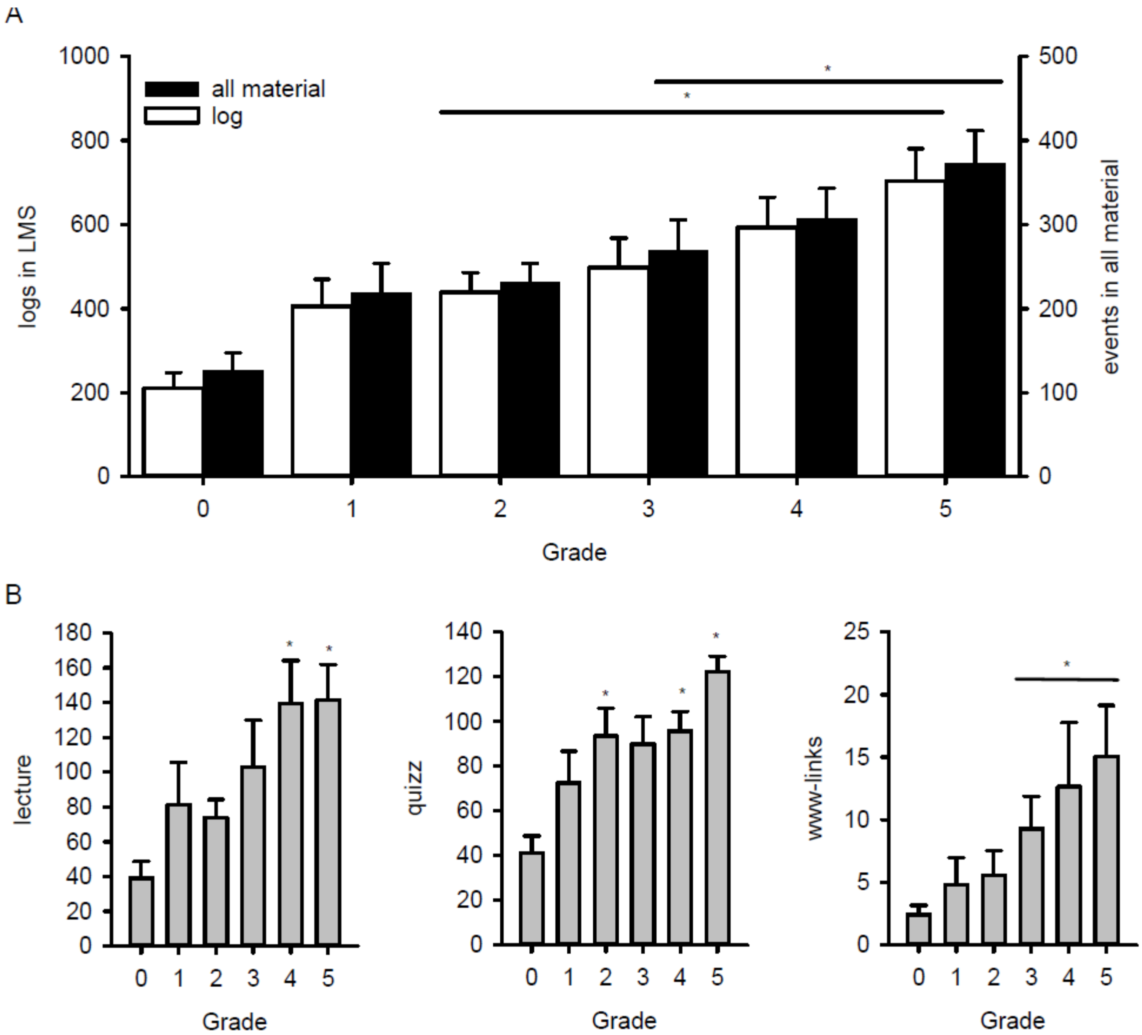

Figure 2: Use of learning components. A) Comparison between learning outcomes, Logs on the LMS during the course and visits on the course pages containing any learning related information. B) Comparison between learning outcomes and visits on video lecture containing course pages (left), quiz containing course pages (middle) and course pages with external links to supplementary material (right). Asterix (*) indicate statistically significant difference between students passing the course with certain grade and the students failing the course (0).

The usefulness of course material was tested by comparing the use of different course material with learning outcome of the student (Fig 2). Studying success had a linear correlation with LMS total use as well as students' visit on learning related content of the course in LMS (Fig 2A). Therefore, harder the student was working with course material, higher grade they got, which indicate a high role of resource use for learning success. Interestingly, the use of LMS and use of course material pages affect grading with different steepness. This small but significant difference is related to log events of starting and ending the learning session, caused by regular study work.

By comparing the use of different course material one can identify content that is in essential role for learning success. Therefore, the number of visits on pages containing video lecture, quiz, or links to external supplementary sources were compared (Fig 2B). Video lectures were in a high use during the course: practically all students visited on lecture containing pages as many times as there are lectures (50). However, students with at least 80 % learning outcome were using more lectures during the course. In comparison, use of quizzes was increased drastically by nearly all students passing the course. Interestingly, LMS contained only 14 sessions of quizzes and students were remaking the same quizzes several times to test their knowledge. Moreover, this seek for a perfect quiz score was done without score affecting the final grading. The course contained 35 links to external sources (video material in English etc.) but these were used only by the students with >70 % learning outcome. Therefore, external material is not helping students to pass the course, but it can motivate students who are looking for more demanding tasks. These results indicate a high importance of self-evaluation and tasks with free allowance to quiz on learning success.

The main difference between blended learning (e.g., flipped classroom) and traditional online course is the interaction between students and teacher. The importance of this learning community on learning success was tested by comparing the learning outcome of students taking part on group work / chat with students participating on the same course without interactive elements. As a member of interactive learning community, students were able to achieve better the learning goals (Table 2). Moreover, interactive students were having drastically increased passing rate: almost 70 % of the students not participating on any social activity were not able to finish and pass the course, whereas drop-out/failure was happening only 4 – 23 % of students who had some social activity. Interestingly, this learning community boost was seen also in students who were only reading other students chat or participating on a single 2 h group work during the course. Therefore, learning community and supportive mechanisms are essential element of successful learning. However, the participation on social events motivated students to study more regularly, using more lecture material and quizzes. Thus, group work was filling its purpose. However, these interactions demonstrate that simple comparison of single elements is unable to identify the best learning strategies.

Table 2
Effects of interactive elements on the course success and study activity

Group work	chat	n	course points (0-150)	pass rate	active learning days	lecture use	quiz use
+	+	24	107,3 ±7,4*	95,8 %*	36,3 ±1,2*	121,2 ±14,8*	101,3 ±8,0*
-	-	29	49,4 ±7,1	31,0 %	20,4 ±2,4	30,3 ±8,9	36,6 ±7,5
+		66	92,4 ±4,3*	77,2 %*	33,8 ±1,5*	106,0 ±10,0*	93,1 ±5,5*
-		32	53,9 ±7,2	34,3 %	20,9 ±2,3	35,7 ±9,5	40,4 ±7,7
	+	27	106,2 ±7,1*	92,5 %*	35,1 ±1,5*	117,6 ±14,1*	98,7 ±8,1*
	-	71	69,7 ±4,6	52,1 %	27,6 ±1,8	69,8 ±9,4	67,2 ±6,1

To reveal the most important parameters affecting learning outcome Best Subset Regression was analyzed by calculating the correlation between learning outcome and all combinations of 1 – 14 parameters (variable). Mallows' Cp had minimum with 5, adjusted Rsquare maximum with 3 and Variance Inflation Factor demonstrated that independency of some parameters in more complex models than 5 parameters was questionable (Table 3). Therefore, five parameter model was describing the learning success with high predictability (R^2=0,645). In comparison, there was no difference between the students taking the course in main and minor study, or students in a degree program of continuous learning (P=0,221), indicating the studying activity have larger role in course success than student's previous studies.

Table 3
Potential key factors of learning

variable	Cp	R^2	Adj R^2	MSer	n(VIF>5)	VIF_{max}	parameters
1	34,038	0,475	0,470	899,251	0	1,000	H
2	8,356	0,589	0,580	712,248	0	1,469	A,H
3	2,247	0,622	**0,610**	661,655	0	1,593	A,H,J
4	-0,299	0,641	0,625	635,823	0	4,788	A,C,H,J
5	**0,609**	0,645	0,626	634,729	**0**	**4,920**	A,C,H,J,K
6	1,619	0,649	0,626	634,368	2	20,274	A,B,D,H,J,K
7	2,846	0,652	0,625	635,630	3	27,164	A-D,H,J,K
8	4,256	0,65	0,624	638,299	2	20,803	A-C,H-L
9	5,683	0,657	0,622	641,162	3	307,804	A-D,H-L
10	7,359	0,659	0,619	646,023	3	30,493	A-E,H-L
11	9,123	0,659	0,616	651,684	3	30,602	A-E,H-M
12	11,014	0,660	0,612	685,483	3	36,261	A-E,G-M
13	13,005	0,660	0,607	666,251	4	39,047	A-M
14	15,000	0,660	0,603	674,238	5	40,126	A-N

Log (A), all material (B), lecture (C), quiz (D), www-links, (E), days in LMS (F), daily use (G), max use (H), active learning days (I), group work (J), chat (K), main study subject (L), starting year (M), sex (N)

A closer look on the model demonstrate that successful learning requires skills of using resources (logs and lecture views), time management skills (max use/all use) as well as use of supportive mechanisms (group work and/or chat) (Table 4). Predictability of some parameters remains still to be elucidated: chat has linear regression with a high risk of detecting non-existing differences (P>0,05). Therefore, relatively small sample size of the study cannot reveal the benefits rarely used chat. Moreover, it was unable to detect the learning though quizzes although their use had a clear correlation with learning outcome (R^2=0,315).

Table 4
Parameters of the model

	r^2	Coef.	s.e.	t	P	VIF
logs	0,435	0,0798	0,0195	4,093	<0,001	4,920
lecture	0,238	-0,137	0,0634	-2,159	0,033	3,945
max use	0,451	-169,195	32,798	-5,159	<0,001	1,546
group work	0,265	3,584	1,235	2,901	0,005	1,270
chat	0,0842	2,043	1,879	1,077	0,284	1,116

To be useful, learning analytics should be able to show the early warning of students with high risk of drop-out/failure. Therefore, students' LMS activity was tracked during the first 3 weeks of the course (Figure 3). Students, whose learning outcome was below 60%, had daily activity on LMS c. 50% of the average activity during the first study week (Fig 3 A). During the second week students passing the course with a low grade (1/5) increased their activity toward average use whereas failing students had LMS activity below 70 % of the average. The differences can be linked to increased participation on group work and regularity of the studies. Students with below 70 % activity were having a high risk of failure/drop-out: passing rate 41,3 and 37,2 % on the day 7 and 14, respectively. In comparison, students with at least 70 % of the average activity were passing the course with 82,7 and 83,6 % passing rate on the same days. Therefore, low activity can be used to predict the students under risks.

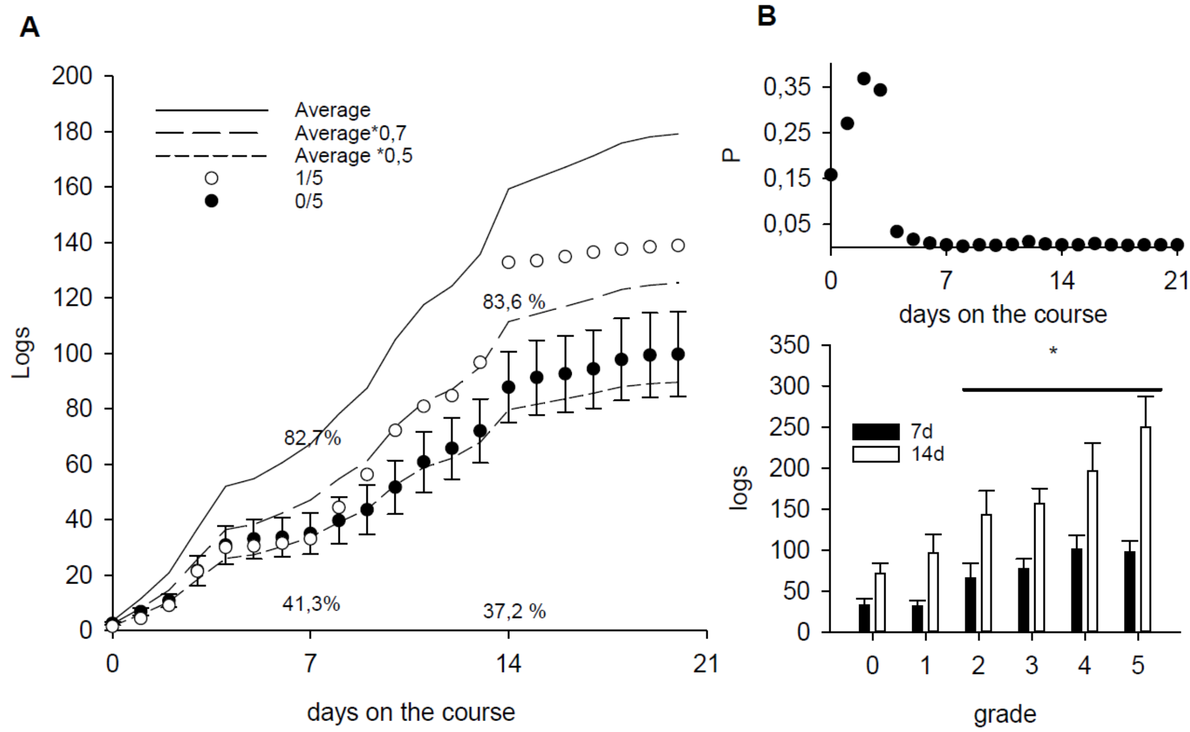

Figure 3: Activity in the first weeks. A) Average cumulative logs, cumulative logs of low-grade passing students (1/5) and failing (0/5). Dashed lines indicate 70% and 50 % of the average activity and numbers show the passing rate of students with above or over 70% of average activity. B) Development of learning activity differences among students with different learning outcomes as P-value (upper) and course activity of students during the first 7 and 14 days. Asterix (*) indicate statistically significant difference between students passing the course with certain grade and the students failing the course (0).

Students with different learning outcome were compared to clarify the speed of developing learning behavioral differences (Fig 3B). After five days on the course there was statistically significant differences (P<0,05 Kruskal-Wallis) on LMS use between students with different learning outcomes. Moreover, after 1st week students achieving at least 60 % of the maximum (i.e. grade 2/5) were using LMS more than students failing the course. Therefore, low LMS activity during the first week of the course predicts a high risk of failure, but differences between failure and low-grade passing are too small for predictions.

6. Discussion

In this study, I have demonstrated that the challenges in time management are in a key role of learning success in online/blended courses (Fig 1, Table 2-3). Therefore, in any courses with summative evaluation, some students will fail the exams as they are not able to study with a steady pace, as has been demonstrated elsewhere [34–36]. This is a strong argument for formative evaluation. In comparison, against my preliminary thoughts, student's static data did not have any effect on the learning success. Therefore, carefully planned course with progressively structured content can help students with any background to reach the learning goals.

I have shown that video lectures in LMS are widely used among the students: almost every student opened each lecture at least once (Fig 2). This was a surprise for me for two reasons. Firstly, based on the pilot flipped classroom courses in our institute, I had assumed lower viewing rate (c. 30 %). Secondly, the lectures were recorded with low-tech video production (screencast). The online lectures reached more students than prior onsite lectures of the same course (c. 30 – 50 %). Therefore, information delivery on the lecture is increased in online environment, for which learners can achieve the learning goals better than in a traditional lecture course.

More detail analysis of the course content revealed the importance of quiz-based self-evaluation and the low importance of external www-material for learning success (Fig 2). Neither of these parameters was explaining the grade in linear form: the use of quizzes was increasing dramatically in low grades whereas external www-material was used only by more advanced students. Therefore, linear models can underestimate or overestimate the value of some course content for learning (Table 2). For the online course developers, these results highlight the value of freely used quizzes for increasing the pass-rate. as has been described elsewhere [45], and limitations of external www-material to motivate students who struggles with the learning content.

This study demonstrated the effectivity of social activity on learning success (Table 2 – 4). Weekly group work meetings and online chat was helping students to maintain the motivation and to reach the learning goals of the course. In several studies students have expressed the difficulty to maintain the motivation and persistence throughout a course [7,34–36], for which online courses without interaction lead easily to a low pass rate. The benefit of social activities exceeded the costs of organizing the group work: ¾ of students participating on group work sessions were able to finish the course successfully whereas only 1/3 of the students without participation on these events were passing the course. Therefore, the presence of teacher during the course is essential and meeting students even in relatively small groups is economically productive.

To be useful for organizations, learning analytics should be able to identify learners at risk, deliver intervention suggestions that work, and be cost-effective [2]. Only then teachers can use learning analytics to help students on their learning progresses. For this purpose, learning analytics can provide information on the key factors of learning success (Table 3 – 4). However, not all these factors are useful in real-time tracking of students' activity [10]. In the current study, proportion of maximal daily use of LMS has strong influence on learning, but naturally, it cannot be measured until the course is over. Therefore, simple uncomplete model can reveal the students at-risk early. For this reason, I tracked the LMS use of the students during the first days of the course and identified statistically significant differences already 4 days after introduction lecture (Figure 3). Similar useful parameters have been described elsewhere [13,14,46]. Average use of course material during the first week of the course is a simple factor to calculate and it can be used to identify students with below certain study rate (e.g., 70 % of average) as they are in high risk of failing or dropping out. It is worth to mention that the current study cannot determinate the exact date or relative LMS activity on which high-risk-students can be identified as the passing rate was increasing linearly with the LMS use even during the first week of the course.

The study demonstrates that by affecting supportive mechanisms, use of resources and time management guidance teacher can help students to reach the learning goals. This support interactive learning models, like CoI, as tools to develop the course [23,25]. Interestingly, the earlier university studies did not help students to achieve learning goals. Especially, students taking the course in their minor studies and studied in the university earlier only onsite had significant difficulties to pass the course. Therefore, support for self-regulated learning and schedule-planning should be offered to students throughout their studies.

6.1. Limitations of the study

Although the current study was tracking all student activity in LMS during the course, it was unable to measure all potentially learning success affecting parameters. During the course in fall 2016 video server of university had closures that lasted several days. These silent periods might have reduced learning activity specially of students with less motivation, or students with large time management challenges. Moreover, the study was not tracking activity outside LMS and interactive elements. Therefore, students can have used podcasts, textbook, and lecture notes for their learning. However, the linear correlation between LMS log information and learning success indicate that the study was using representative sample of parameters for reliable conclusions.

The predictability of the learning related factor model identified in this study remains to be elucidated. The models generated for a specific course can have limited usability in general [47]. More research is needed to find the best practices, how teachers can modulate online and blended teaching, help students to fully participate early on the course studies, and increase the pass rate of the courses without reduction of the learning goals.

Acknowledgements

I would like to thank Prof. Erkko Sointu for his valuable comments on this study project.

References

[1] C. Klein, R. Hess, Using Learning Analytics to Improve Student Learning Outcomes Assessment, in: J. Lester, C. Klein, A. Johri, H. Rangwala (Eds.), Learning Analytics in Higher Education: Current Innovations, Future Potential and Practical Applications, Routledge, New York, 2019: pp. 140–159.

[2] B. Rienties, S. Cross, Z. Zdrahal, Implementing a Learning Analytics Intervention and Evaluation: What works?, in: B. Daniel (Ed.), Big Data and Learning Analytics in Higher Education. Current Theory and Practice, Springer Int., Switzerland, 2017: pp. 147–166.

[3] L. Vygotsky, The problem of the cultural development of the child. The Pedagogical Seminary and Journal of Genetic Psychology. 36 (1929) 415–434.

[4] M. Stains, J. Harshman, M.K. Barker, S. v. Chasteen, R. Cole, S.E. DeChenne-Peters, M.K. Eagan, J.M. Esson, J.K. Knight, F.A. Laski, M. Levis-Fitzgerald, C.J. Lee, S.M. Lo, L.M. McDonnell, T.A. McKay, N. Michelotti, A. Musgrove, M.S. Palmer, K.M. Plank, T.M. Rodela, E.R. Sanders, N.G. Schimpf, P.M. Schulte, M.K. Smith, M. Stetzer, B. van Valkenburgh, E. Vinson, L.K. Weir, P.J. Wendel, L.B. Wheeler, A.M. Young, Anatomy of STEM teaching in North American universities, Science. 359 (2018) 1468–1470. https://doi.org/10.1126/SCIENCE.AAP8892.

[5] R. Ubell, Going Online. Perspectives on Digital Learning, Taylor & Francis, New York, 2017.

[6] L. Lockyer, E. Heathcote, S. Dawson, Informing Pedagogical Action: Aligning Learning Analytics With Learning Design, American Behavioral Scientist. 57 (2013) 1439–1459. https://doi.org/10.1177/0002764213479367.

[7] A. Hadwin, J. Nesbit, D. Jamieson-Noel, J. Code, P. Winne, Examining trace data to explore self-regulated learning, Metacognition Learning. 2 (2007) 107–124.

[8] J. You, Examining the effect of academic procrastination on achievement using LMS data in e-learning, Educational Technology & Society. 18 (2015) 124–134.

[9] S. Graf, C. Ives, N. Rahman, A. Ferri, A Tool for Accessing and Analysing Students' Behaviour Data in Learning Systems, in: LAK2011 Proceedings of the Conference on Learning Analytics & Knowledge, Banff, Alberta, Canada, 2011.

[10] L. Morris, C. Finnegan, S. Wu, Tracking student behavior, persistence, and achievement in online courses, The Internet and Higher Education. 8 (2005) 221–231.

[11] P. Prinsloo, S. Slade, Big Data, Higher Education and Learning Analytics: Beyond Justice, Toward an Ethics of Care, in: Daniel BK (Ed.), Big Data and Learning Analytics in Higher Education, Springer, Switzerland, 2017: pp. 109–124.

[12] L. Macfadyen, S. Dawson, Mining LMS data to develop an "early warning system" for educators: A proof of concept, Computers & Education. 54 (2010) 588–599.

[13] G. Johnson, Student alienation, academic achievement, and WebCT use, Educational Technology & Society. 8 (2005) 179–189.

[14] A. Wang, M. Newlin, Predictors of web-student performance: The role of self-efficacy and reasons for taking an on-line class, Computers in Human Behavior. 18 (2002) 151–163.

[15] Ö. Vural, The Impact of a Question-Embedded Video-based Learning Tool on E-learning, Educational Sciences: Theory and Practivr. 13 (2013) 1315–1323. www.edam.com.tr/estp.

[16] A.C. Butler, J.D. Karpicke, H.L. Roediger, Correcting a Metacognitive Error: Feedback Increases Retention of Low-Confidence Correct Responses, Journal of Experimental Psychology: Learning Memory and Cognition. 34 (2008) 918–928. https://doi.org/10.1037/0278-7393.34.4.918.

[17] J. Cranney, M. Ahn, R. McKinnon, S. Morris, K. Watts, The testing effect, collaborative learning, and retrieval-induced facilitation in a classroom setting, European Journal of Cognitive Psychology. 21 (2009) 919–940. https://doi.org/10.1080/09541440802413505.
[18] M. Vojdanoska, J. Cranney, B.R. Newell, The testing effect: The role of feedback and collaboration in a tertiary classroom setting, Applied Cognitive Psychology. 24 (2010) 1183–1195. https://doi.org/10.1002/acp.1630.
[19] A. Butler, Repeated Testing Produces Superior Transfer of Learning Relative to Repeated Studying, Journal of Experimental Psychology: Learning, Memory, and Cognition. (2010). https://doi.org/10.1037/a0019902.supp.
[20] D.R. Garrison, T. Anderson, W. Archer, Critical Thinking, Cognitive Presence, and Computer Conferencing in Distance Education, The Internet and Higher Education. 2 (2000) 87–105.
[21] N. Croft, A. Dalton, M. Grant, Overcoming Isolation in Distance Learning: Building a Learning Community through Time and Space, Journal for Education in the Built Environment. 5 (2010) 27–64.
[22] E. McElrath, K. McDowell, Pedagogical Strategies for Building Community in Graduate Level Distance Education Courses, MERLOT Journal of Online Learning and Teaching. 4 (2008).
[23] R. Fasse, J. Humbert, R. Rappold, R.O. Learning, Rocherster institute of technology: Analyzing student success, Journal of Asynchronous Learning Networks. 13 (2009) 37–48. http://www.rit.edu/emcs/ptgrad/online/.
[24] C. Hostetter, M. Busch, Measuring Up Online: The Relationship between Social Presence and Student Learning Satisfaction, Journal of Scholarship of Teaching and Learning. 6 (2006) 1–12.
[25] P. Shea, K. Swan, C.S. Li, A. Pickett, Developing learning community in online asynchronous college courses: The role of teaching presence, Online Learning Journal. 9 (2019) 59–82. https://doi.org/10.24059/OLJ.V9I4.1779.
[26] M. Marmon, The value of social presence in developing student satisfaction and learning outcomes in online environments, in: R.D. Wright (Ed.), Student-Teacher Interaction in Online Learning Environments, IGI Global, Hershey, 2015: pp. 120–132.
[27] S.K. Mitchell, M. Shepard, Building social presence through engaging online instructional strategies, in: R.D. Wright (Ed.), Student-Teacher Interaction in Online Learning Environments, IGI Global, Hershey, 2015: pp. 133–156.
[28] R.A. Bjork, J. Dunlosky, N. Kornell, Self-regulated learning: Beliefs, techniques, and illusions, Annual Review of Psychology. 64 (2013) 417–444. https://doi.org/10.1146/annurev-psych-113011-143823.
[29] D.H. Schunk, B.J.Z. Schunk, D.H. Zimmerman, Social origins of self-regulatory competence, Educational Psychologist. 32 (1997) 195–208. http://www.tandf.co.uk/journals/.
[30] G. Rakes, K. Dunn, The impact of online graduate students' motivation and self-regulation on academic procrastination, Journal of Interactive Online Learning. 9 (2010) 78–93.
[31] P. Sun, R. Tsai, G. Finger, Y. Chen, D. Yeh, What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction, Computers & Education. 50 (2008) 1183–1202.
[32] J.W. You, Identifying significant indicators using LMS data to predict course achievement in online learning, Internet and Higher Education. 29 (2016) 23–30. https://doi.org/10.1016/j.iheduc.2015.11.003.
[33] E. Yukselturk, S. Bulut, Predictors for student success in an online course, Educational Technology & Society. 10 (2007) 71–83.
[34] G. Elvers, D. Polzella, K. Graetz, Procrastination in online courses: Performance and attitudinal differences, Teaching of Psychology. 30 (2003) 159–162.
[35] Y. Levy, M. Ramin, A study of online exams procrastination using data analytics techniques, Interdisciplinary Journal of E-Learning and Learning Objects. 8 (2012) 97–113.
[36] N. Michinov, S. Brunot, O. le Bohec, J. Juhel, M. Delaval, Procrastination, participation, and performance in online learning environments, Computers and Education. 56 (2011) 243–252. https://doi.org/10.1016/j.compedu.2010.07.025.
[37] C. Asarta, J. Schmidt, Access patterns of online materials in a blended course, Decision Sciences Journal of Innovative Education. 11 (2013) 107–123.

[38] M.J. Lage, G.J. Platt, M. Treglia, Inverting the classroom: A gateway to creating an inclusive learning environment, Journal of Economic Education. 31 (2000) 30–43. https://doi.org/10.1080/00220480009596759.
[39] J. O'Flaherty, C. Phillips, The use of flipped classrooms in higher education: A scoping review, Internet and Higher Education. 25 (2015) 85–95. https://doi.org/10.1016/j.iheduc.2015.02.002.
[40] R. Talbert, Flipped classroom A guide for Higher Education Faculty, Stylus publishing, LCC, Sterling, 2017: pp 1–264.
[41] J. Broadbent, S. Sharman, E. Panadero, M. Fuller-Tyszkiewicz, How does self-regulated learning influence formative assessment and summative grade? Comparing online and blended learners, Internet and Higher Education. 50 (2021). https://doi.org/10.1016/j.iheduc.2021.100805.
[42] M. Hyypiä, E. Sointu, L. Hirsto, T. Valtonen, Key components of learning environments in creating a positive flipped classroom course experience, International Journal of Learning, Teaching and Educational Research. 18 (2019) 61–86. https://doi.org/10.26803/IJLTER.18.13.4.
[43] L. Hyppönen, L. Hirsto, E. Sointu, Perspectives on university students' self-regulated learning, task-avoidance, time management and achievement in a flipped classroom context, International Journal of Learning, Teaching and Educational Research. 18 (2019) 87–106. https://doi.org/10.26803/IJLTER.18.13.5.
[44] S. Chatterjee, J. Simonoff, Handbook of regression analysis, Wiley, New York, NY, 2013.
[45] J. Dunlosky, K.A. Rawson, E.J. Marsh, M.J. Nathan, D.T. Willingham, Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology, Psychological Science in the Public Interest, Supplement. 14 (2013) 4–58. https://doi.org/10.1177/1529100612453266.
[46] M. Brown, Learning Analytics: Moving from Concept to Practice, Educause. (2012) 1–5.
[47] S. Knight, A. Gibson, A. Shibani, Implementing learning analytics for learning impact: Taking tools to task, Internet and Higher Education. 45 (2020). https://doi.org/10.1016/j.iheduc.2020.100729.