Practice and Repetition during Exam Preparation in Blended Learning Courses: Correlations with Learning Results

Monika Andergassen, Felix Mödritscher, and Gustaf Neumann

Information Systems and New Media

Vienna University of Economics and Business (WU)

Austria

monika.andergassen@wu.ac.at

ABSTRACT: Learner-centric research on factors influencing learning results has focused, among other things, on student characteristics, demographic data, and usage patterns in learning management systems (LMSs). This paper complements the existing research by investigating potential correlations between learning results and LMS usage during exam preparation, focusing on practice and repetition. Based on 250 million log-file entries used to analyze student interactions within specific courses and overall in the LMS, results show positive, albeit modest, correlations between usage variables and final exam grades. Regarding practice, the number of learning days and the number of days between the first and the last learning sessions correlate better than the coverage of different learning materials. The findings for repetition indicate that it is more beneficial to transfer learning to new tasks than to repeat the same items many times. The study not only looks at single usage variables but also examines the distribution of the descriptive and dependent variables and uses visualization techniques and quantiles to deal with outliers. This paper describes the largest empirical study of learner interactions in blended learning courses conducted so far (at least according to the authors' knowledge) and including techniques for processing and analyzing large datasets about LMS usage.

KEYWORDS: Learning analytics, web usage mining, learning management systems, blended learning courses, practice, repetition, correlation analysis, regression analysis

1 INTRODUCTION

In the context of higher education, learning management systems (LMSs) are particularly important for coping with the phenomenon of mass education (Johnson et al., 2012). Consequently, LMS platforms that provide online and blended learning courses collect large amounts of interaction data from students (cf. Ferguson, 2012). Learning Analytics (LA) attempts to exploit user-generated data through Business Intelligence techniques in order to support different stakeholders, from students and teachers to LMS developers and service providers, and to predict learning performance in educational settings (Siemens et al., 2011). Additionally, large sets of learning-related data can be of interest for research purposes, i.e., to understand how students learn and to advance an LMS.

Several universities in Austria teach massive cohorts of students in the first semesters of their study programs. At the Vienna University of Economics and Business (WU), for instance, some courses in the Bachelor study programs have 1,000 students or more per class. As a result, the LMS has become an essential instrument for managing large groups of students and for providing multiple-choice tests to assess student performance and learning outcomes. This paper examines correlations between learning results and LMS usage variables in relation to the practice and repetition of course material. It draws on previous research focusing on seasonal effects in LMS usage data (Andergassen, Neumann, & Mödritscher, 2013) and dependencies between LMS usage in terms of page hits, user sessions and learning results (Mödritscher, Neumann, & Andergassen, 2013). In order to do this, the log-files of the Learn@WU LMS are analyzed and compared with other data sources, namely, the final exam grades achieved by students of three selected courses. Thanks to the considerable number of users and the intense use of the LMS, large datasets were available for this study.

We proceed in the following way. First, we examine the theoretical foundations of LMS usage, with a particular focus on practice and repetition. We also indicate the limitations of the existing research. Next, we point out the aims and objectives of our research, and describe the research methodology for the empirical study, in which we apply Web Usage Mining techniques and statistics to show dependencies between LMS usage variables and learning results. The "Results and Discussion" section summarizes the most relevant findings gained from this study. The last section concludes the paper and gives an outlook for future research.

2 FOUNDATIONS OF RESEARCH ON LMS USAGE AND LIMITATIONS

2.1 LMS Usage in Learning Analytics Research

In recent years, research has first proposed and then pushed the development of Learning Analytics, which seems to reveal a promising insight into students' technology usage behaviour within educational processes. Furthermore, it also provides the basis for improving the situation for various stakeholders, from learners and teachers to service providers, developers, and organizations (Siemens et al., 2011). The latter, in particular, have started to investigate analyzing LMS usage systematically because of data-driven research.

According to Chatti, Dyckhoff, Schroeder, and Thüs (2012), research in the field of Learning Analytics principally focuses on the adaptation of learning environments (40%), monitoring and analysis (33%), assessment and feedback (13%) and the prediction of student performance (12%). Other objectives, such as reflection or supporting competence development are (still) underrepresented. For LMS technology, LA seems to be useful for identifying didactic or technical flaws (e.g., inadequate or incorrect task assignments in courses or usability problems with the system); predicting trends (e.g., the learning outcomes in connection with different factors); detecting patterns in LMS usage data (e.g., at-

risk vs. highly gifted students); and supporting learning and competence development (e.g., by regularly providing performance indicators and visual elements).

Research so far has also included topics such as course persistence; for instance, correlations between demographic data and course persistence (Doherty, 2006). Furthermore, cumulations of activities throughout semester tertiles has served to identify different types of persistence in a course, such as "low-extent-users," "late users" and "online-quitters" (Hershkovitz & Nachmias, 2011). Similarly, changes in LMS activity have been used to identify students at risk of course attrition (Wolff, Zdrahal, Nikolov, & Pantucek, 2013). Whitmer's (2012) study of learning results investigates links between LMS usage variables and student characteristics, i.e., the demographic data of the highly diverse student population and their grades. Other approaches include the comparison of web users and mobile users according to their usage behaviour within the LMS, or the identification of seasonal effects by analyzing activity patterns in log-files (Andergassen et al., 2013). With regard to didactic design, research has considered topics such as the impact of online discussions on learning (Khan, Clear, & Sajadi, 2012; Wise, Zhao, & Hausknecht, 2013). Early research results by Mödritscher, Neumann, et al. (2013), whose study explored correlations between the number of different learning days and the overall learning time, indicate that practice and repetition play an important role in determining final exam performance.

2.2 Practice and Repetition

The analysis of LMS usage highly depends on the didactical model implemented in the online and blended learning courses (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Mödritscher, Andergassen, Law, & García-Barrios, 2013). Courses with LMS platforms are often used to distribute learning materials and to provide self-assessment tests. Consequently, it is possible to track students' repetition and practicing of the course contents within the LMS itself through LA approaches.

Once information has been received by the working memory, practice is important in order to establish this data in the long-term memory (Willingham, 2004). In his meta-study, Cotton (1989) examined practice in relation to time, comparing time factors with achievements. She investigated differences between the time for learning allocated by teachers and the real times of student engagement, including dead time as well as learning times above and below the relevant experience levels. Her findings were that the allocated time showed a slightly positive relationship with learning results (grades), while the time-on-task had a positive relationship and the times of learning a strong one.

Another large body of research reports on positive correlations between spacing effects and learning (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Dunlosky et al., 2013; Thalheimer, 2006; Vlach & Sandhofer, 2012; Wells & Hagman, 1989). Thalheimer (2006) reviews research on spacing effects and related learning factors and finds, among other things, that repetition is effective for learning, spaced repetition is generally more efficient than non-spaced repetition and spacing is beneficial for long-time retention. Similar findings are reported by Cepeda et al. (2008) and Vlach and Sandhofer (2012).

Regarding repetition, Rawson and Dunlosky (2011) differentiate between recalling and relearning. With a focus on the mnemonic strategy of retrieval practice, they examined the effects of initial learning criterion (i.e., a required number of correct recalls) and relearning (i.e., the repetition of test items after an extended time period) on the durability and efficiency of learning. Here, it was shown that learning efficiency is influenced by the number of trials with correct recalls, while durability depends on the rate of relearning. Increasing the initial learning criterion showed a strong effect on retention in the absence of repetition. This effect diminished markedly through relearning. Relearning was thus shown to be more costly in terms of time but also reduced the number of attempts to reach the initial learning criterion.

2.3 Discussion and Limitations of Former Findings on LMS Usage

Summing up this section, early research results from the field of LA have yielded promising insights into LMS usage. However, more research is needed about LMS usage in general and about practice and repetition in particular. Firstly, most existing research either deals with small datasets or small samples (Andergassen et al., 2013; Wise et al., 2013), investigates LMS activities without differentiating the activity types (Hershkovitz & Nachmias, 2011), or is restricted to one course only (Whitmer, 2012). Although Whitmer (2012) gives evidence that technology usage can be more useful than learner characteristics to predict student performance, it should be pointed out that these findings are based on one online course with 377 students and that the LMS usage of students beyond this course has not been considered at all. Moreover, the study only examines a limited number of demographic variables (e.g., being from a racial/ethnic under-represented minority); five LMS usage categories (i.e., administration, assessment, content activity, engagement activity, overall course activity); and only one usage indicator (i.e., the number of hits).

Secondly, although research has pointed to the importance of practice and repetition in learning, most existing empirical results for this topic go back to the time before widespread LMS usage (Wells & Hagman, 1989); are based on rather small setups (Mödritscher, Andergassen, et al., 2013; Vlach & Sandhofer, 2012); or incorporate only small sets of usage variables (Cepeda et al., 2008; Cotton, 1989; Mödritscher, Neumann, et al., 2013). In addition, an important aspect is the focus on a real-world setting given by the log-files of an LMS platform, whereas most of the research mentioned above uses experimental settings. While the outcome of practicing an exercise item — as measured by Rawson and Dunlosky (2011) — cannot be logged by a regular log-file entry, our log-file analysis gives insights into students' real-world learning behaviour. The research presented in this paper thus aims to complement existing research by tackling these issues.

2.4 Aims and Objectives

In this study, therefore, we address the following: a) a substantial set of LMS usage variables; b) all

activities of learners in a platform not restricted to one course; and c) several courses that have high numbers of participants and can be assigned to different knowledge domains. Drawing on the studies mentioned above, usage data makes it possible to measure activities both within and outside a specific course; interactions with specific LMS applications such as quizzes and exercises; and the repetition of actions, spacing of actions, and learning time within the LMS. This article reports on our research efforts on analyzing LMS usage and on an empirical study that we have conducted at our university. In particular, we investigate potential correlations between student e-learning usage patterns and final exam grades. The following research questions are addressed in this paper:

1. How are final grades related to practice and repetition in LMS usage while preparing for an exam?
2. How does preparing for a specific exam relate to the overall usage of an LMS in the exam preparation period?
3. Does exam preparation vary between different course domains, and which commonalities can be found?

The empirical study is based on the LMS of the Vienna University of Economics and Business (WU), called Learn@WU. The WU ranks among the largest business and economics universities worldwide and is one of the largest universities in Austria. The Learn@WU system covers all university courses (about 5,000 courses per year), which are predominantly held in a blended learning mode. One of the traditional problems of the (public) university is the heterogeneous knowledge of first-year students, which results in different learning paces among students with different backgrounds. This was one of the reasons for a strong emphasis on self-assessment facilities in the LMS in the first place. To improve the students' throughput and to decrease the period of study for the most gifted students, the university introduced half-semester terms.

Thus, in their first year of study, students can pass a semester-long course in half a semester with the help of e-learning. If a student fails the exam, he/she can enroll in the course again in the second half of the semester. Failure rates of 50% or more are not unusual in these exams. The university offers three exam weeks per semester for all beginner courses: one at the beginning, one in the middle and one at the end of the semester. The exam weeks are preceded by a so-called "exam preparation week." In these weeks, the LMS usage rates are the highest in the whole year, with up to 3.8 million page views per day. Students solve up to 600,000 self-assessment exercises per day (Mödritscher, Neumann, et al., 2013). With these figures, Learn@WU is one of the most intensely used e-learning systems worldwide.

Against this background, our engagement in LA aims to gain a better understanding of the student learning processes in this study phase. The overall goal from a mid- and long-term view is to improve the courses in terms of learning processes and didactic models based on the empirical data from the LMS.

# 3 RESEARCH METHODOLOGY

3.1 Context and Data Sources

The research was conducted based on data from the 14 days of the exam preparation week and the exam week in November 2012. The raw data comprised LMS server log-files containing 250 million entries, which means an average of about 17 million hits per day.

The research layout focuses on the following: 1) a comparison of LMS usage in three blended courses from different knowledge domains, comprising elementary business, law, and IT topics; and 2) a comparison of LMS usage within such a course with the usage in all other areas of the LMS. The selected courses are from the beginning phase of the Bachelor program. Usually, students attend these courses in parallel with a set of additional courses. From the three courses, we analyzed the overall LMS usage of all students who attended the final exam of these courses, using data retrieved from a period of the 14 days immediately before the exam. Due to the high number of self-test learning materials, the usage of the LMS during this time span is the highest in the semester. All three courses are half-semester courses concluding with a final exam, which consists of a paper and pencil multiple-choice test. The course C1 comprised n=883 participants, C2 comprised n=389 participants, and C3 comprised n=578 participants. Typically, the failure rate in these courses is relatively high (39.8% in C1, 45.8% in C2, and 46.0% in C3).

3.2 Web Usage Mining

In order to analyze student LMS usage, we followed the Web Usage Mining process suggested by Srivastava, Cooley, Deshpande, and Tan (2000), which consists of three major steps.

The first step, ***data preprocessing***, deals with the selection and transformation of data. As displayed in Figure 1, we selected all entries from the raw data of the students that participated in at least one of the three courses examined in our study (i.e., the log-files from the 14 days before the exam, which were available in a slightly extended Combined Log format). This led to three datasets containing 2.3 million entries (course C1), 1.2 million entries (course C2), and 1.6 million entries (course C3), respectively.
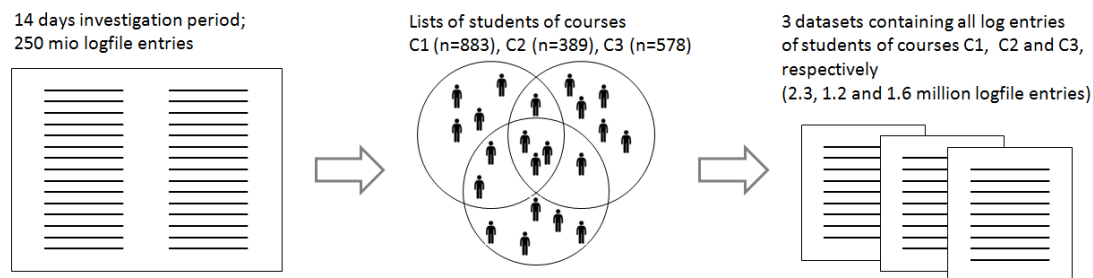
Figure 1: Generation of datasets for all log-file entries of students attending at least one of the three courses
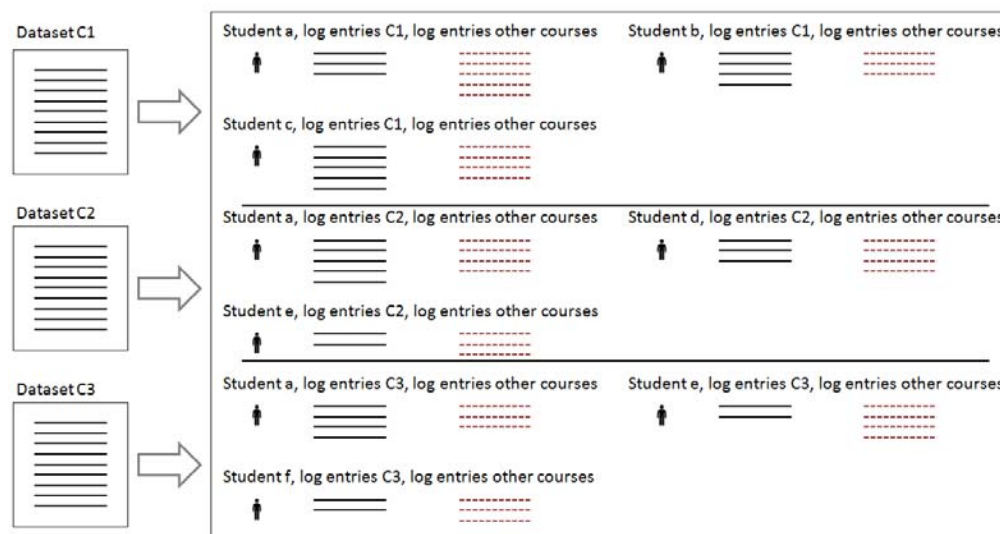
We then sorted the log-file entries according to each student's actions within the specific course and outside this specific course (Figure 2). Overall, the distribution of student activity in the LMS can be summarized as follows: 788,517 activities (i.e., page views) in C1 vs. 386,913 activities in 166 other courses; 229,260 activities in C2 vs. 444,238 activities in 124 other courses; 731,694 activities in C3 vs. 456,849 activities in 139 other courses. Next, the log and exam data were connected, and all data was anonymized.

The second step, ***pattern discovery***, aims to extract patterns of LMS usage and to generate condensed data structures, such as usage variables and descriptive statistics. In our case, we a) extracted the user sessions from the log-files and b) defined and calculated usage variables for practice and repetition within the targeted course and within the LMS for each student. The third step, ***pattern analysis***, applies analysis techniques to the condensed data on usage patterns. For our study, we applied inferential statistical methods from correlation and regression analysis, combined with visual representations of the analysis results.

In all three phases of this Web Usage Mining process, we proceeded in an explorative way. To be precise, we followed an iterative process, whereby the definition of usage variables led to inferential statistics about their correlation with final grades, and the results of that led to the definition of new sets of variables or the need to apply other methods of analysis, and so on.

Figure 2: Sorting of log-file entries according to each student's actions within the specific course and outside the specific course

3.3 Iterative Selection and Refinement of Usage Variables and Analysis Methods

We started with a set of variables that operationalizes practice and repetition. The variable selection

was approached from a time-related and a learning-material-related perspective. Regarding **practice**, the user sessions made it possible to calculate learning times by adding together the times between the time stamps of log-file entries in the sessions. Furthermore, the time stamps allowed us to detect the number of different learning days. Two kinds of learning materials for measuring student progress in practicing were investigated, namely, self-assessment exercises and sample exam questions. These materials, which are the most intensely used on Learn@WU, are particularly intended for self-assessment and preparation for the final exam.

The usage of each interactive exercise and exam question is captured in the log-files by means of a unique identifier. The amount of usage can therefore be calculated as a further indicator for practice. Through the HTTP methods GET and POST in the log-file entries, it is possible to distinguish between activities related to the instruction or solution pages of an exercise.

Regarding **repetition**, the repetitive use of these learning materials was counted, again based on their unique identifiers. Time gaps between the repeated solving of self-assessment exercises and sample exam questions were also investigated. Since the log-files and user sessions contain information about courses through the URLs, it was possible to compare usage within and beyond the three individual courses. While the literature suggests having longer gaps between learning sessions (i.e., from one day to several weeks) to foster long-term retention (cf. Cepeda et al., 2008; Dunlosky et al., 2013), this study focuses on short-term retention in order to pass exams in a Bachelor program. Time gaps were calculated based on days with an expected range from 0 to 14 days. A more fine-grained analysis would have been possible but was not conducted in this paper.

As a next step, descriptive statistics such as frequencies, means, and medians were produced and correlation coefficients were calculated. Some variables were complemented with or evolved into more fine-granular variables. Over several iterations, we defined and refined a set of 83 LMS usage variables and calculated them for all students of the three courses. Due to the limited space, this paper concentrates on those variables that revealed the highest correlations with final grades. Table 1 gives an overview of these variables. The usage variables in lines 2 to 10 are defined as follows:

- **topic | other:** The students usually attend many parallel courses in each semester. They might not distribute their learning evenly among the courses, and learning intensely for one course might negatively affect the learning in other courses. To study such possible displacement competition, the identifier {topic|other} refers to the measurement of all usage variables regarding student activities within a course under investigation (*topic*) and overall LMS usage and thus student activities in other courses (*other*). For instance, the variable "duration" was calculated for each student's visiting duration within the course C1 (or C2 and C3, respectively), and the visiting duration in other courses. Thus, two variables were calculated for the duration time: *topic.duration* and *other.duration*.

- **excs | exam:** The identifier {excs|exam} indicates the investigated learning materials, namely, self-assessment exercises and sample exam questions. For instance, the variable "different" was calculated for each student's number of unique solved self-assessment exercises (*excs*) and sample exam questions (*exam*).

Table 1: Overview of the LMS usage variables extracted from raw data

Dependent Variable	Description
Points, grade	Points and grade achieved in the final exam
Variables as indicators for practice	
{topic|other}.duration	Absolute duration (hours) of each student's activities. This variable is calculated by adding up the time intervals between each of a student's learning activities (page hit).
{topic|other}.days	Absolute number of different days each student used the LMS. The variable *topic.days* counts each day that contains at least one activity by the student within the course. The variable *other.days counts e*ach day that contains at least one activity by the student in other courses. Thus, the range varies from 0 (no activities) to 14 days (investigation time span).
{topic|other}.dayspan	Time span (days) between first and last activity. This variable complements the variable {topic|other}.days by indicating whether a student has distributed his learning time evenly throughout the days, or, for example, has only started to prepare shortly before the exam.
topic.{excs|exam}_different	Absolute number of unique interactive solved exercises and sample exam questions. The three investigated courses contain up to about 1,500 self-assessment exercises and sample exam questions. The variable measures how many different items a student has solved.
topic.{excs|exam}_coverage	Similar to *_different*, but as a percentage.
topic.{excs|exam}_avg_considertime	Average time used to consider exercises and exam questions (i.e., work towards the solution). The investigation distinguishes between page views of the exercise problem and page views of the exercise solution. The variable *_avg_considertime* is the ratio of the total time spent on viewing a problems page versus the sum of all the exercises/exam questions solved by the student.
Variables as indicators for repetition	
topic.{excs|exam}_repeat_coverage	Percentage of unique self-assessment exercises and sample exam questions that were solved more than once
topic.{excs|exam}_repeat_factor	Ratio of total solved self-assessment exercises and sample exam questions versus different solved items
topic.{excs|exam}_avg_repeat_gap	Average time differences between repetitions in solving self-assessment exercises and sample exam questions

The variables *.duration*, *.days* and *.dayspan* were calculated both for in-course (topic) and out-of-course (other) activities. The variables *_difference*, *_coverage* and *_avg_considertime* were only

calculated for in-course activities. The results of these show some differences between self-assessment exercises (*excs*) and sample exam questions (*exam*), since the variables relating to the entire LMS usage besides the specific course (*other.**) are close to null in these cases. As a result, we conducted the final pattern analysis with a set of 12 variables to investigate practice in LMS usage and a set of six variables to investigate repetition.

The correlation of the variables with the final exam grades was calculated using Pearson correlations. Moreover, the significance levels were determined as follows. Null-hypotheses about each independent variable's correlation with the final grades were formulated. The significance level of 5% was corrected using the conservative Bonferroni method, thus leading to a new significance level of 0.42% for the 12 variables representing practice and 0.83% for the 6 variables representing repetition. Since the correlation coefficients only give a limited impression of the nature of the dependencies, bag plots and scatter plots were created to get a more detailed picture. Bag plots are bivariate extensions of box plots, and make it possible to analyze the distribution of two variables in relation to each other. Scatter plots display the values of two variables on the axis of a Cartesian coordinate system and are useful in identifying dependencies between different variables in a visual manner.

Scatter plots can be prone to outliers, for instance because of the smoothing of the non-parametric regression curve. Therefore, quantile box plots, a visualization technique for graphically depicting groups of numerical data through five-number summaries (i.e., the median, the lower and upper quartiles, as well as the sample minimum and maximum) for different quantiles of a usage variable, were applied. Furthermore, quantile regression comprises a method for regression analysis that aims to estimate either the conditional median or other quantiles of the response variable, and is thus more robust against outliers in the response measurements (Koenker, 2008) that we were facing in our dataset. For this kind of regression analysis, we used the R package "quantreg" (Koenker, 2013). We created quantile regression plots, which consist of box plots for the dependent variable (points), and contrasted the quantile regression and ordinary least square (OLS) estimates of selected covariates.

Finally, we observed a particularity in the data of course C1. The exam for C1 consists of two partial exams. If a student does not reach a minimum of points in either partial exam, he is assigned 0 points for the entire exam. This caused a highly uneven distribution of final points in our plots, giving the illusion that many students were not able to gain any points at the exam. Therefore, we filtered out students with 0 points, which reduced the number of participants in C1 from n=883 to n=786.

3.4 Research Ethics

Research ethics are important within LA research, including issues such as learner rights and data ownership. However, still lacking is an ethical framework that defines, for instance, the rights of learners in relation to their data, including opting out of the analytics record and giving informed consent for data usage to researchers (Ferguson, 2012).

To counter this, Slade and Prinsloo (2013) suggest six principles towards an ethical framework. They include the following points: 1) LA should function as moral practice; 2) students should be seen as agents and thus as collaborators in developing their learning; 3) data collected about performance should be seen as temporal, dynamic constructs and thus should have an expiry date; 4) student success should be seen as a multidimensional phenomenon far beyond learning analytics metrics; 5) transparency about data usage should be offered by the university; and 6) higher education needs to use LA better to understand and develop outcomes for students.

In our study, the following measures were taken to ensure the rights of the students. Users of the Learn@WU platform were informed about data storage and data analysis by the university upon registration. Any analysis of student-related data occurs in an anonymized form to protect student privacy as far as possible. As in the current case, the unique identifiers assigned to the dataset through k-anonymization make a re-identification of the students theoretically possible. Data that allow the back-tracing of individual students without their informed consent is not published at all.

Regarding the sixth principle of Slade and Prinsloo, our aim in LA research is to advance the LMS in terms of quality and effectiveness of teaching and learning. In order to do this, student-related data may need to be collected, but only to the extent absolutely necessary.

4 RESULTS AND DISCUSSION

4.1 Practice

4.1.1 Learning time, learning days, and dayspan correlate positively, albeit modestly, with final exam points

There is a positive relationship between the total duration, the learning days, and the dayspan in the course on the LMS and final points. Table 2 summarizes these variables for all three courses, including the Pearson correlation coefficients. The table shows that all correlation coefficients are based on a good sample size (i.e., the number of all participants in the courses) and most of them also have good significance levels when considered individually, i.e., p-values smaller than 0.0042, indicating that the correlation is unlikely to result from random chance. This also holds when considering the 12 practice-related variables and null-hypotheses interdependently by applying the Bonferroni correction method. Tables 2 and 3 highlight the p-values above the corrected significance level of 0.42%.

The correlation is the highest in course C2, with Pearson correlation coefficients of 0.41 for *topic.days* and 0.40 for *topic.dayspan*. The numbers show indications, albeit on a modest level, that the more different days a student prepares for the exam on average, the better are his/her results. For C1 and C3, these correlations are lower (C1: 0.29 and 0.25; C3: 0.24 and 0.21).

58

It can also be observed that student behaviour regarding learning days and dayspan varies between the courses. While in C1 the students study on 11 days with a median dayspan of 12, and thus almost every day of the investigation period, the learning days and dayspan in C3 (6 and 10 days, respectively) and C2 (5 and 9 days, respectively) are lower. The weaker correlations in course C1 indicate that, for this course, other factors play a more important role for the final grades than regular online learning.

Table 2: Mean, Median, and Pearson correlation coefficients for the variables *.duration, *.days and *.dayspan; all three courses (Bonferroni-corrected significance level: 0.0042)

Duration (hours)	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Topic	*Other*	*Topic*	*Other*	*Topic*	*Other*
Mean	25.05	5.11	6.00	18.39	12.62	13.57
Median	24.17	1.62	3.74	15.72	8.54	8.20
Pearson corr. coeff.	**0.3006**	0.0889	0.2672	0.0328	0.2379	0.0848
	(p<2.2e-16)	(**p=0.0127**)	(p=8.7e-8)	(**p=0.5195**)	(p=7.0e-9)	(**p=0.0417**)
Learning days	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Topic	*Other*	*Topic*	*Other*	*Topic*	*Other*
Mean	10.30	8.88	5.21	9.47	6.57	9.05
Median	11.0	10.0	5.0	11.0	6.0	10.0
Pearson corr. coeff.	0.2861	0.2063	**0.4119**	0.1504	0.2390	0.2400
	(p=4.4e-16)	(p=5.3e-9)	(p<2.2e-16)	(p=0.0029)	(p=5.9e-9)	(p=5.1e-9)
	C1 (n=786)		C2 (n=389)		C3 (n=576)	
Dayspan	*Topic*	*Other*	*Topic*	*Other*	*Topic*	*Other*
Mean	10.90	10.30	7.29	10.40	8.18	9.91
Median	12.0	11.0	9.0	12.0	10.0	11.0
Pearson corr. coeff.	0.2476	0.2056	**0.4037**	0.2038	0.2105	0.2221
	(p=1.9e-12)	(p=6.0e-9)	(p<2.2e-16)	(p=5.1e-5)	(p=3.3e-7)	(p=6.8e-8)

4.1.2 Displacement competition between courses is not observed

To study the displacement competition of online learning activities, we calculated the correlation of *other.duration* (online time spent by a student preparing for an exam outside the topic of the exam), *other.days* and *other.dayspan* with the points achieved in the exam. As shown in Table 2, the correlations with learning results are either very low (*other.duration*) or rather positive (*other.days*, *other.dayspan*). Moreover, the learning duration outside the examined courses shows little significance, as the *p*-values for *other.duration* are above the Bonferroni-corrected significance level of 0.0042 in all courses. The positive correlation coefficients indicate no displacement competition between online learning for one course and online activities in other courses.

4.1.3 Coverage of exercises and exam questions correlates positively, albeit modestly, with final exam points

Regarding learning materials, the correlation between the usage of self-assessment exercises and sample exam questions and the achieved results was investigated. In absolute numbers, course C1 contains 272 self-assessment exercises and 184 sample exam questions; course C2 contains 593 exercises and 105 sample exam questions; and course C3 contains 1,368 exercises and 120 sample exam questions. This is represented by the usage variable *_different of Table 3.

Table 3: Mean, Median, and Pearson correlation coefficients for the variables topic.*_different, topic.*_coverage and topic.*_avg_considertime; all three courses (Bonferroni-corrected significance level: 0.0042)

Different	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Excs	*Exam*	*Excs*	*Exam*	*Excs*	*Exam*
Mean	101.0/272	63.3/184	93.8/593	43.5/105	333/1368	41.9/120
Median	93/272	47/184	2/593	34/105	193/1368	10/120
Pearson corr. coeff.	0.2489 (p=1.5e-12)	0.3132 (p<2.2e-16)	0.2390 (p=1.9e-6)	0.2243 (p=7.9e-6)	0.3388 (p<2.2e-16)	0.2610 (p=1.9e-10)
Coverage	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Excs	*Exam*	*Excs*	*Exam*	*Excs*	*Exam*
Mean (percent)	37.03	34.39	15.53	41.41	23.91	34.89
Median (percent)	34.19	25.54	0.33	32.38	13.85	8.33
Pearson corr. coeff.	0.2489 (p=1.5e-12)	0.3132 (p<2.2e-16)	0.2390 (p=1.9e-6)	0.2243 (p=7.9e-6)	**0.3388** (p<2.2e-16)	0.2610 (p=1.9e-10)
Avg. considertime	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Excs	*Exam*	*Excs*	*Exam*	*Excs*	*Exam*
Mean (seconds)	173.0	166.0	35.9	38.7	37.8	31.4
Median (seconds)	171.0	178.0	15.0	37.9	36.3	31.1
Pearson corr. coeff.	0.0846 (**p=0.0177**)	**0.3742** (p<2.2e-16)	0.06013 (**p=0.2368**)	0.1480 (p=0.0034)	0.1007 (**p=0.0155**)	0.1667 (p=5.6e-5)

C3 has the largest set of self-assessment exercises, which explains the rather small median percentage. The usage variable *_coverage expresses the percentage of learning materials (exercises and sample exam questions) solved by the students when preparing for the final exam within the two weeks of investigation. Table 3 shows positive correlations between topic.excs_different and topic.exam_different (and *_coverage, respectively) and final exam points. The correlation is strongest between topic.excs_coverage and final points in the course C3. The variable topic.excs_avg_considertime is less

significant with *p*-values of up to 0.2368. The Bonferroni correction only has effects on the average considertime for exercises in all courses, whereby the correlation is very low in these cases.

To obtain a better understanding of the learning material coverage, we first provide an overview of the medians (*topic.excs_coverage*, *topic.exam_coverage*) and achieved points. Such dependencies can be plotted as bag plots. The bag plots in Figure 3 show the variable *topic.excs_coverage* on the *x*-axis, and number of points attained in the exam on the *y*-axis. In course C1, a student has to achieve 71 points to pass the test, in C2 30 points and in C3 24 points. The bag plots mark the medium 50% of values, which are included in the box plot between the first and third quartile. The outer (light blue) area corresponds to the length of the whiskers. The points located outside the blue area are outliers and plotted as individual points.
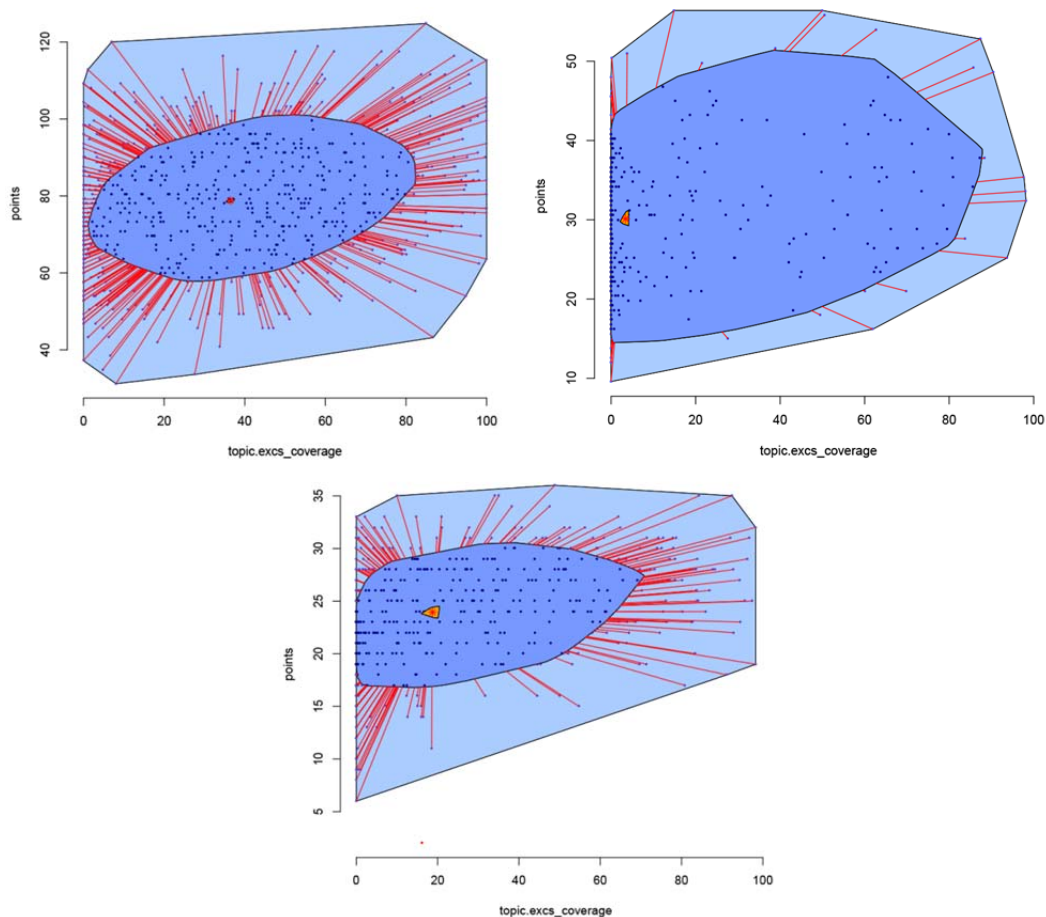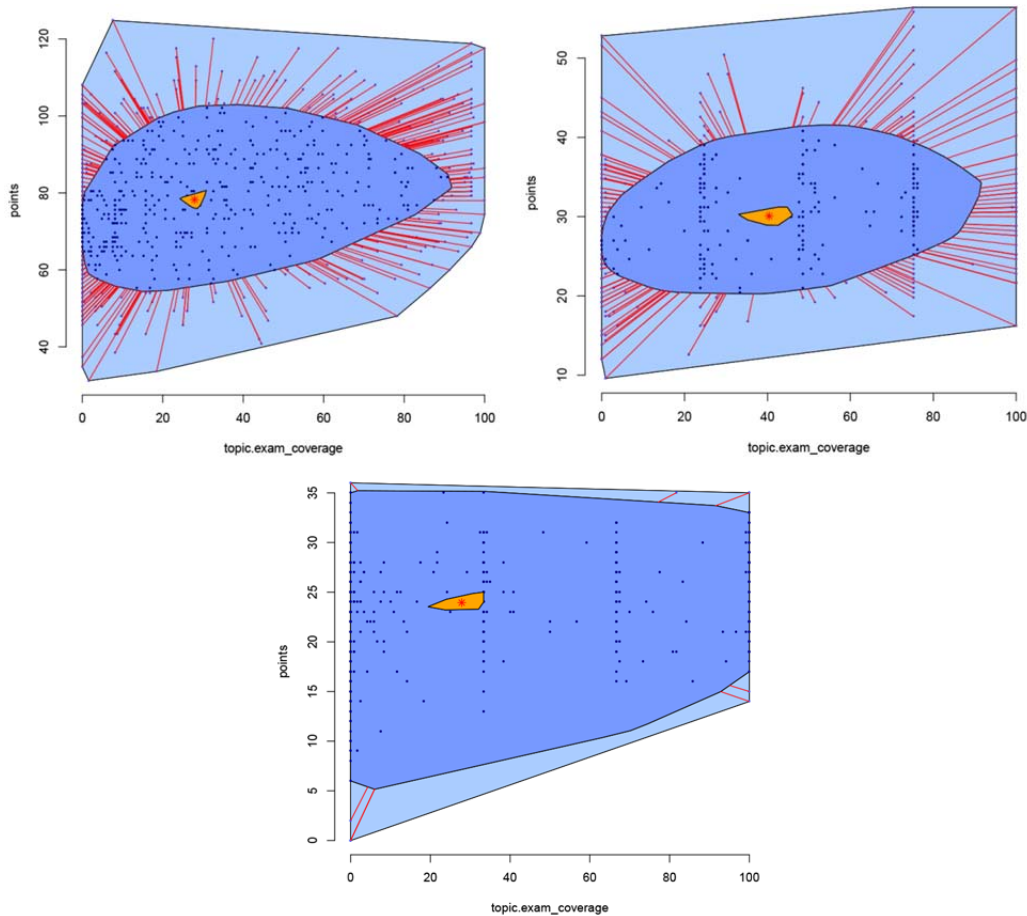

Figure 3: Bag plots with topic.excs_coverage and final points for C1 (top left), C2 (top right), and C3 (bottom)

The bag plots show that the median student solves about 40% of self-assessment exercises and gains about 80 points in course C1 (Figure 3, top left). The range between the first and third quartile of students in the inner (dark blue) area shows that 50% of students get between 60 and 100 points,

solving between about 3% and 80% of exercises on the topic in the LMS. In the light blue area at the bottom left, there are the students who have not solved many exercises but also have not gained many points. The top right of the plot depicts those students who have solved many exercises with good results. At the bottom right, the students who have solved high percentages of exercises for the topic but have not passed the exam are situated. Some students also gained good results without substantial use of the LMS (top left).

Figure 4: Bag plots with topic.exam_coverage and final points for C1 (top left), C2 (top right), and C3 (bottom)

The median of *topic.excs_coverage* is very low for C2 compared to C1 (0.33% vs. 34.19%), although the correlation coefficients are similar (0.24 vs. 0.25). For all three courses, the 50% population bubble reaches about 80% coverage of exercises, despite the fact that C3 offers about 5 times more exercises than C1 (Table 3). Thus, while the average coverage of solved exercises in course C2 is much smaller than in C1 and C3, the coverage of sample exam questions is much higher. Figure 4 shows this relationship. Furthermore, many students solve 100% of the sample exam questions in course C2, in contrast to C1 and C3. Solving sample exam questions also correlates higher with the final points in C2 than solving exercises.

4.1.4 Oscillation effect and regression of exercise coverage and final exam points

The scatter plots of Figure 5 visualize the relationship between the usage variable *topic.excs_coverage* and final points in more detail. The scatter plots can be read as follows. As above, the *x*-axis shows the values of the explanatory variable, which in our case is *topic.excs_coverage*. The *y*-axis shows the values of the outcome variable, which in our case are the final points. The green straight line is the linear regression line (OLS). In our study, this rises in each plot and thus marks positive relationships. The red solid line is the non-parametric regression line, and the red dotted lines are the nonparametric regression lines for the first and third quantiles. The scatter plots also contain box plots to show dependencies between the explanatory and outcome variables. For instance, the plot of C1 (top left) can be read as the average student solving about 38% of self-assessment exercises and gaining 78 points.
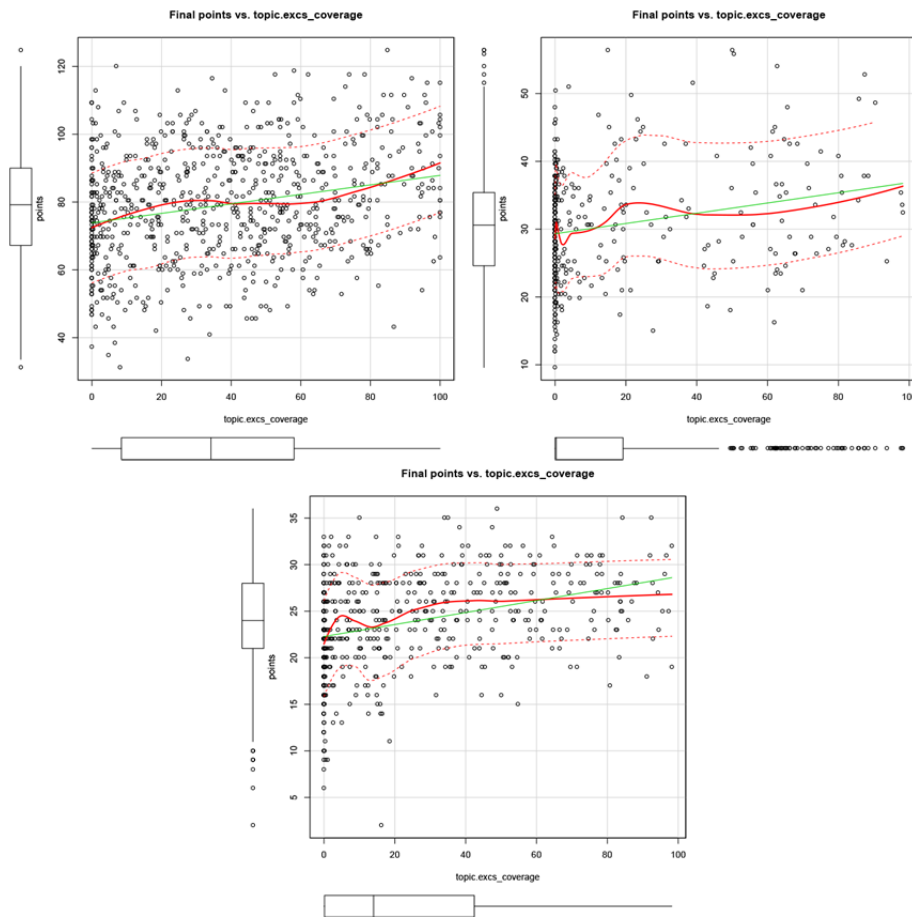

Figure 5: Scatter plots with topic.excs_coverage and final points for C1 (top left), C2 (top right), and C3 (bottom)

Course C1 shows a nearly linear distribution, while some oscillation effect is present in courses C2 and C3. Moreover, it becomes apparent that in courses C2 and C3, many students do not solve a significant fraction of the self-assessment exercises during exam preparation. Nevertheless, their final points cover

the entire range from very low to very high scores, and some candidates achieve good results without significant LMS usage. The non-parametric regression curve of C3 (bottom) becomes horizontal towards the end, indicating that after a certain number of exercises, saturation is reached where solving more exercises does not correlate with better results. The right hand side of the scatter plot is in a sparsely populated area, so there is no strong significance. However, the diminishing gain from additional exercises is clearly visible. Furthermore, the regression curve is oscillating in courses C2 and C3.

These observations might be caused by either outlier or smoothing effects in the scatter plot splines. Therefore, we investigated the phenomena further by introducing quantile box plots and quantile regression plots. The quantile box plots and quantile regression were made for the explanatory variable *topic.excs_coverage* and final points in course C2. The quantile box plot (Figure 6, left) shows the quantiles of *topic.excs_coverage* in relation to the final points. It only contains the values of the students who solved at least one interactive exercise (n=222/389 students). In the quantile box plots, the usage variable *topic.excs_coverage* is divided into 10 quantiles. As before, we observe a significant oscillation in the box plot of C2, this time at the 0.3 quantile. The students in the high quantiles (e.g., at 0.9) of *topic.excs_coverage* do not score significantly higher than those of the low quantiles.

While the analysis with the multiple box plots used quantiles for more robust results on the explanatory variable *topic.excs_coverage*, quantile regression (Figure 6, right) uses the quantiles of the outcome variable to study the effects of solved exercises for successful and unsuccessful students. Table 4 lists the OLS regression coefficients, the quantile regression coefficient at the 0.1, 0.5, and 0.9 quantiles for the explanatory variable *topic.excs_coverage* and the outcome variable for all three courses. The values of the quantile regression coefficients differ significantly from the OLS regression coefficient in course C2. In the 0.9 quantile, each additional percentage of solved exercises would lead to 0.12 more points in a literal interpretation.
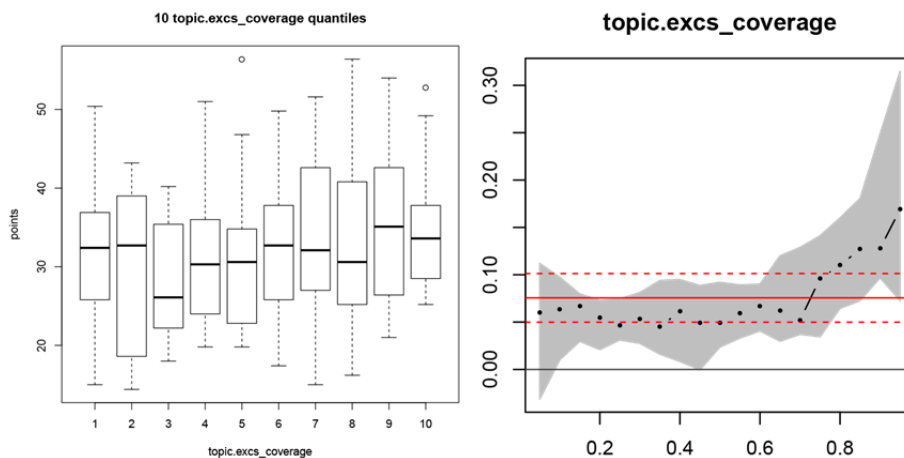
Figure 6: Quantile box plot (left; n=222) and quantile regression (right; n=389) for the explanatory variable topic.excs_coverage and the depending variable final points; course C2

Table 4: OLS regression and quantile regression coefficients for the usage variable topic.excs_coverage

Total points	OLS regression (coefficient β, std. error σ)	Quant. regression at 0.1 quantile	Quant. regression at 0.5 quantile	Quant. regression at 0.9 quantile
C1 topic.excs_coverage (percent)	β=0.14, σ=0.0195 (p=1.5e-12)	0.1682	0.1204	0.1321
C2 topic.excs_coverage (percent)	β=0.0756, σ=0.0156 (p=1.9e-6)	0.0636	**0.0491**	**0.1277**
C3 topic.excs_coverage (percent)	β=0.0668, σ=0.0077 (p<2.0e-16)	0.0914	0.0582	0.0438

The quantile regression plot of Figure 6 (right) can be read as follows: The quantiles of the outcome variable (points achieved) are listed on the x-axis and the coefficient magnitudes on the y-axis. The red solid and dotted lines mark the OLS coefficient and the OLS confidence interval. The OLS coefficient is equal throughout all the quantiles and shows the correlation between *topic.excs_coverage* and final points (one additional percent of solved exercises relates on average to 0.14 points for C2). The black dotted line marks the connection between quantile regression coefficients; the grey area is the confidence interval. If the black line is outside of the OLS confidence interval, then it can be concluded that there is a strong deviation between the OLS regression and the quantile regression.

It can be clearly observed in Figure 6 that the higher quantiles (students with higher grades) benefit strongly from solving self-assessment exercises. This deviation only starts at the 0.7 quantile. For the higher quantiles, the quantile regression line is steeper than the OLS regression line, indicating that for better performing students, the variable *topic.excs_coverage* had a more positive correlation. This positive deviation from the OLS regression is also reflected in the quantile regression coefficients of Table 4.

Summing up, learning through practicing self-assessment exercises does positively correlate to final points. However, there are differences in the three courses. In course C2, where the exercises demand abstract thinking, we can observe oscillation effects. Furthermore, many students do not practice through solving self-assessment exercises. Particularly in course C2, solving sample exam questions is a more widely used practice among students than solving self-assessment exercises. To gain the highest grades, however, it is beneficial to solve high percentages of self-assessment exercises.

The time taken to solve exercises differs depending on the course domain
Finally, we investigated how long students take to solve exercises and sample exam questions. As the numbers in Table 3 show, the usage variables *topic.*_avg_considertime* vary strongly between the courses. While the times are similar in C2 and C3 and range between 31 and 39 seconds on average, they are much higher in C1, at about 2.5 to 3 minutes. This can be explained by the nature of the

subjects. While materials in C1 include many items where calculations have to be made, the exercises and sample exam questions in C3 are mainly text-based.

4.2 Repetition

While the focus of the previous section was on results regarding practice, this section presents results regarding indicators for repetition in LMS usage.

4.2.1 Repeated solving of exercises and exam questions does not strongly correlate with final grades
With a focus on the learning materials, Table 5 summarizes the mean values and Pearson correlation coefficients of variables including the repeat coverage, the factor of the repetition, and the average repeat gap. Again, we have corrected the significance level for the six repetition-related variables according to the Bonferroni method, which has led to the insight that the correlation coefficients for the exam-related variables in C2 and for *exam_repeat_avggap* in C3 are less significant.

Table 5: Mean, Median, and Pearson correlation coefficients for the variables topic.*_repeat_coverage, topic.*_repeat_factor and topic.*_repeat_avggap; all three courses (Bonferroni-corrected significance level: 0.0083)

	C1 (n=786)		C2 (n=389)		C3 (n=576)	
Repeat coverage	*Excs*	*Exam*	*Excs*	*Exam*	*Excs*	*Exam*
Mean	12.45	16.28	6.47	17.59	5.92	10.58
Median	5.88	8.15	0.00	6.67	1.33	0.83
Pearson corr. coeff.	0.1494 (p=2.6e-5)	0.1960 (p=3.0e-8)	0.1837 (p=2.7e-4)	0.0560 (**p=0.2704**)	0.2121 (p=2.7e-7)	0.1725 (p=3.1e-5)
Repeat factor	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Excs	*Exam*	*Excs*	*Exam*	*Excs*	*Exam*
Mean	1.27	1.10	0.81	1.15	0.98	0.68
Median	1.22	1.07	1.00	1.00	1.12	1.00
Pearson corr. coeff.	0.1334 (p=1.8e-4)	0.1164 (p=0.0011)	0.1757 (p=5.0e-4)	0.0875 (**p=0.0849**)	0.2242 (p=5.1e-8)	0.2375 (p=7.5e-9)
Repeat avggap	C1 (n=786)		C2 (n=389)		C3 (n=576)	
	Excs	*Exam*	*Excs*	*Exam*	*Excs*	*Exam*
Mean(hours)	28.45	25.76	7.28	21.60	18.70	16.80
Median(hours)	14.12	6.66	0.00	0.00	2.60	0.00
Pearson corr. coeff.	0.1926 (p=5.3e-8)	0.2075 (p=4.3e-9)	0.1410 (p=0.0053)	0.0812 (**p=0.1097**)	0.1955 (p=2.2e-6)	0.1049 (**p=0.0116**)

The numbers suggest that on average, not much repetition is done when preparing for the exams. This can also be noted in the bag plots in Figure 7 (*topic.excs_repeat_factor*). The repeat coverage shows that, in courses C2 and C3, repeating exercises correlates more strongly with grades than repeating sample exam questions. As these two courses have more exercises, the repeat coverage is lower than in C1. Furthermore, the average repeat gap for exercises in C2 is much shorter than for C1 or C3. It becomes apparent that "coverage"-variables (see the previous section) have higher correlations with the final points than the "repeat"-variables. This indicates that solving many different exercises and exam questions is more effective than repeating exercises and exam questions.

Therefore, although the variables reveal differences in the example usages between the courses, they do not serve as good predictors for results in an overall analysis. Below, parametric regression and quantile regression analysis is used once again for a more detailed view.
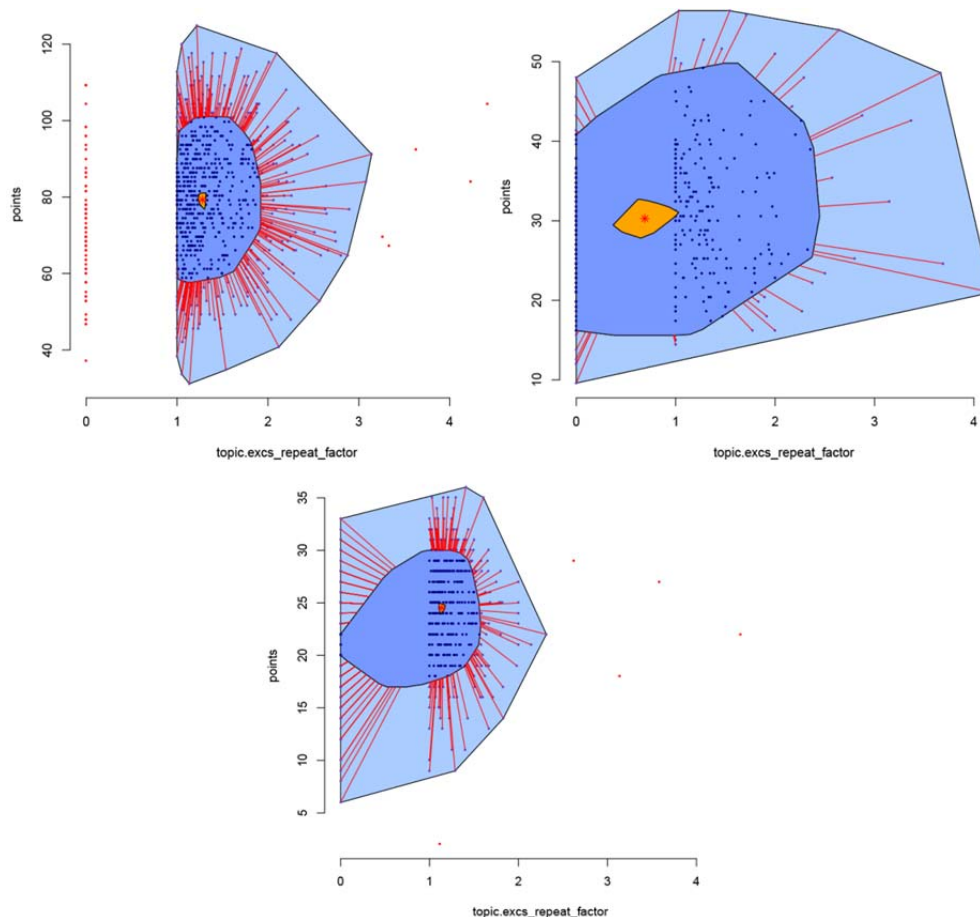
Figure 7: Bag plots with topic.excs_repeat_factor and final points for C1 (top left), C2 (top right) and C3 (bottom)

*4.2.2 Repeat factor: A "bend" in repeated solving of exercises*

The dependencies between the repeat *topic.excs_repeat_factor* and final points can be plotted as bag plots (Figure 7). In course C1, the inter-quartile 50% of students (dark blue area) has a repeat factor of

between 1 and 1.9 and final points ranging from 58 to 100 points. The inter-quartile bags of courses C2 and C3 also contain items with a repeat factor of 0, meaning that these students did not solve any exercises. This observation is similar to the one made in the previous section. Indeed, the variable *topic.excs_repeat_factor* builds upon the coverage of exercises (*topic.excs_repeat_coverage*) attempted by the students.

There are, however, some characteristics in the distributions. The scatter plots of Figure 8 show that repeat factors of 1 (C1) or slightly above (C2 and C3) relate well to high final points. A repeat factor of 1 in this context indicates that, on average, a student solved exercises once; a repeat factor of 1.3 indicates that, on average, a student solved exercises 1.3 times, etc. After these factors, the curves drop in all three courses.
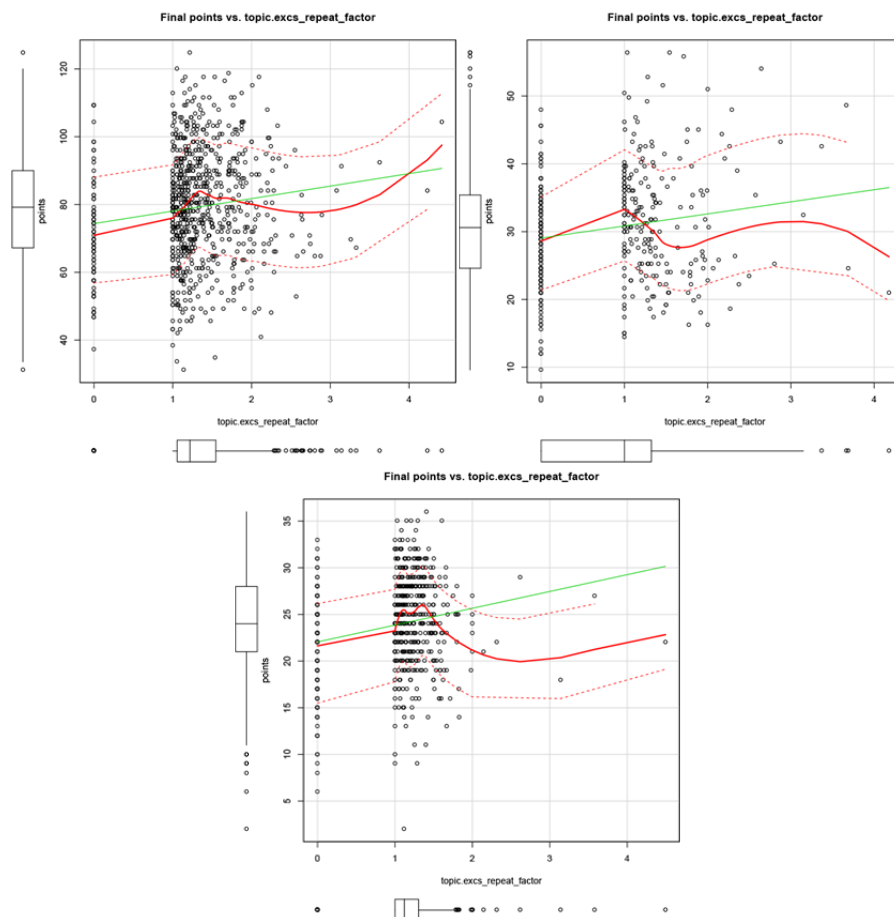
Figure 8: Scatter plots with topic.excs_repeat_factor and final points for C1 (top left), C2 (top right), and C3 (bottom)

The quantile box plot of Figure 9 (left) further analyzes the characteristics of the scatter plots and bag plots, using course C2 as an example. The quantile box plot only contains the values of those students whose *topic.excs_repeat_factor* is greater than 0, and thus represents 222 values. Some similar

characteristics to the scatter plot (Figure 8 top right) can be observed. There is a small rise from the first to the second quantile. Then the median drops, with a small exception in the sixth quantile. The best result is reached in the second quantile with a repeat factor of between 1.01 and 1.07, and in the tenth quantile with a repeat factor of between 2.07 and 4.18. It becomes apparent that the quantile box plot differs slightly from the scatter plot, which seems to be less robust against outliers (particularly on the far right).

Finally, Figure 9 (right) shows the quantile regression plot for the final point quantiles and the explanatory variable *topic.excs_repeat_factor*. As can be seen in Table 6 and Figure 9, the quantile regression deviates strongly from the OLS regression. In other words, students with low grades benefit much less from repetition than the students in the high quantiles.
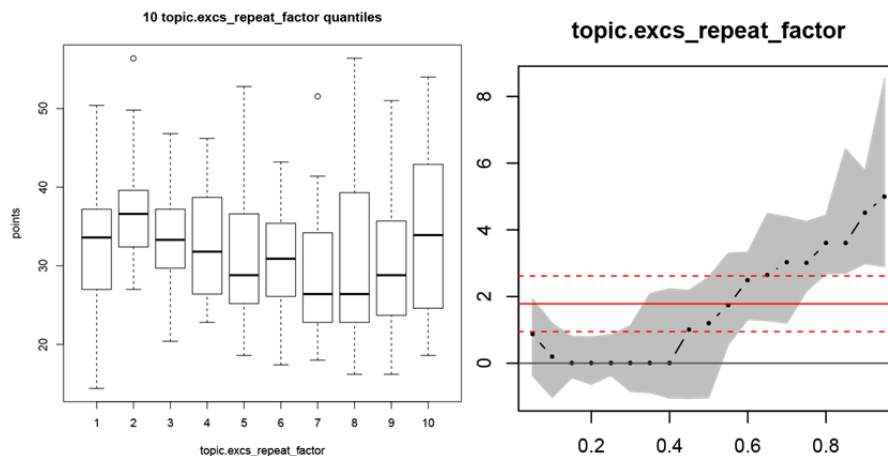
Figure 9: Quantile box plot (left n=222) and quantile regression (right; n=389) for the explanatory variable topic.excs_repeat_factor and the depending variable final points; course C2

Table 6: OLS regression and quantile regression coefficients for the usage variable topic.excs_repeat_factor

Total points	OLS regression (coefficient β, std. error σ)	Quant. regression at 0.1 quantile	Quant. regression at 0.5 quantile	Quant. regression at 0.9 quantile
C1 topic.excs_repeat_factor	β=3.687, σ=0.978 (p=1.8e-4)	4.218	3.600	2.057
C2 topic.excs_repeat_factor	β=1.783, σ=0.508 (p=5.0e-4)	**0.189**	**1.200**	**4.500**
C3 topic.excs_repeat_factor	β=2.009, σ=0.364 (p=5.1e-8)	2.000	1.122	1.956

4.2.3 Due to low levels of repeating activity, correlations regarding spacing between repetitions and final grades are of limited interest

The variables *topic.excs_repeat_avggap* and *topic.exam_repeat_avggap* as indicators of spacing between repeating an exercise or sample exam question have limited informative value regarding final

grades. As analyzed above, the coverage of repeated learning materials and the numbers of repetitions are quite low, and therefore any variables deduced from there cover small proportions of items too. Furthermore, Table 6 shows low Pearson correlation coefficients.

5 CONCLUSION AND FUTURE WORK

This paper analyzed dependencies between practice and repetition in an LMS and final exam results. The LMS usage was analyzed based on 250 million log-file entries collected from the Learn@WU platform over a time span of 14 days during the exam preparation phase. Three courses of different knowledge domains were investigated, and variables were calculated for usage both within and outside the selected courses. Regarding practice, we found positive, albeit modest, correlations between the amount of the learning time, learning days, dayspan of learning, and the coverage of self-assessment exercises and sample exam questions and the final exam points in each course. Repeated solving of exercises and exam questions does not correlate strongly with grades; the best results are achieved when the items are practiced once or twice only. The repeating activity in our test data is rather low. The correlations between final grades and repeat factor, as well as the spacing between repetitions of learning material items, should therefore be treated with caution.

Our analysis covers the exam preparation time of blended learning courses. We know from questionnaire data that students in the courses under investigation used the LMS for more than 60% of their learning time, but we have no detailed analysis of their offline activities. With regard to the literature, the findings are in line with Cotton (1989) in terms of learning time having a positive, albeit modest, relationship with student achievement. However, the correlations are below the effect of time-on-task activity (d=0.38) reported by Hattie (2009). Furthermore, while the positive relationship found between distributing learning over several days (learning days and dayspan) and final grades confirms earlier findings by Vlach and Sandhofer (2012), the relationship is again smaller than Hattie's (2009) spaced versus mass practice (d=0.71). The deviating numbers might be because our study only considers online activities and thus no in-class time and offline learning. As well, the study only considers the brief period of 14 days of exam preparation.

Regarding repetition, our findings only partially correlate with the literature. Where Wells and Hagman (1989) and Thalheimer (2006), for instance, find that, generally, repetition is necessary to achieve proficiency, our findings only show modest correlations between final grades and the repetition of exercises and sample exam questions.

On the other hand, Wells and Hagman also argue that, depending on the goal of learning, multiple repetitions are helpful for the learning of a specific task, while fewer specific task repetitions in combination with a broader task variety promote transfer to other tasks. Our findings indicate that it is more beneficial to transfer learning to new tasks than to repeat items often. Indeed, the exercises and sample exam questions in the investigated courses cover a broad variety of tasks and therefore promote

the transfer of learning to different tasks. The correlations are higher between *topic.different* (as a representation of task variety) and final grades than between *topic.{excs|exam}_repeat_factor* (as a representation of the repeated solving of exercises) and final grades in all three courses. However, when compared to Wells and Hagman's results, both the task variety and the repetition of items only show modest correlations with final exam grades. The low correlations could therefore be an indicator of the need to advance the didactic design of the LMS by providing more possibilities for teachers to design spaced repetition events. Examples include the writing of summaries and elaborative interrogation (cf. Dunlosky et al., 2013).

The overall learning material-usage data also did not correlate as strongly with the student's grades as we would have expected. Among others, one reason might be that students have different learning strategies and start with a heterogeneous array of knowledge. Consequently, in order to understand this better, it is not sufficient to look at a single variable. In contrast, it is necessary to examine the distribution of the descriptive and dependent variables and to use visualization techniques and quantiles to deal with outliers. Based on scatter plots, non-parametric regression, and quantile regression, we were able to identify saturation effects and even negative effects (oscillation effects) on higher numbers of online learning activities.

Regarding the relationship between preparing for a specific exam and the overall usage of an LMS, no displacement effects were found. Thus, a student who spends time in the course classes and in other courses is still likely to perform well in the final exam. Indeed, the correlations were small and rather positive between LMS usage within a specific course and final grades and between LMS usage in other courses and final grades. However, there are some differences between the courses. Oscillation effects in the scatter plots of the self-assessment exercise coverage in the three courses indicate that sometimes students achieve good exam grades with little LMS usage. The prevalent use of either self-assessment exercises or sample exam questions also differs among the courses. Moreover, the time spent on the individual exercises varies strongly between the courses, which might be due to the different course domains.

There are certain limits to the interpretation of the observed variables. The detailed explanation of these variables requires deeper knowledge about the didactic design of the courses. The data is most meaningful for the professors teaching these courses, because they can correlate observed behaviour in class with the planned online learning activities. In contrast, our data does not cover all the students' learning activities, but only their online ones. It is therefore impossible to draw conclusions about the total learning effort and strategies of the students, or about the quality or suitability of the learning design or materials. Another limitation is that in order to investigate dependencies between LMS usage and final exam grades, only students who received points in the final exam were considered, and thus the activities of other users (e.g., the lecturer or students without a final grade) were filtered out at this stage.

Future work will head into two directions. Although the three blended courses examined in this paper cover a variety of knowledge domains, we plan to extend the analysis over a full semester and across other courses. It would also be interesting to perform a longitudinal study to measure the stability of the results over multiple exams and semesters and to observe the impact of spacing over longer periods. Here, we would face the challenge of having to process significantly more data but we also see the potential for learning more about usage variables and pedagogical factors that correlate to learning in positive or negative ways.

ACKNOWLEDGEMENTS

REFERENCES

Andergassen, M., Neumann, G., & Mödritscher, F. (2013). The four seasons: Identification of seasonal effects in LMS usage data. Presented at the *Alpine Rendez-Vous 2013: Workshop on Data Analysis and Interpretation for Learning Environments (DAILE '13)*, Villard-de-Lans, France.

Cepeda, N.J., Vul, E., Rohrer, D., Wixted, J.T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, *19*(11), 1095–1102. doi:10.1111/j.1467-9280.2008.02209.x

Chatti, M.A., Dyckhoff, A.L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5/6), 318–331. doi:10.1504/IJTEL .2012.051815

Cotton, K. (1989). *Educational time factors* (No. Close-Up #8). Northwest Regional Educational Laboratory. Retrieved from http://educationnorthwest.org/webfm_send/564

Doherty, W. (2006). An analysis of multiple factors affecting retention in web-based community college courses. *The Internet and Higher Education.* doi:10.1016/j.iheduc.2006.08.004

Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. doi:10.1177/1529100612453266

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304. doi:10.1504/IJTEL.2012.051816

Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. In L.H. Meyer, S. Davidson, H. Anderson, P.M. Fletcher, P.M. Johnston, & M. Rees (Eds.), *Tertiary assessment & higher education student outcomes: Policy, practice & research* (pp. 259–275). Wellington, New Zealand: Ako Aotearoa.

Hershkovitz, A., & Nachmias, R. (2011). Online persistence in higher education web-supported courses. *The Internet and Higher Education*, *14*(2), 98–106. doi:10.1016/j.iheduc.2010.08.001

Johnson, J., Shum, S.B., Willis, A., Bishop, S., Zamenopoulos, T., Swithenby, S., … Helbing, D. (2012). The FuturICT education accelerator. *The European Physical Journal Special Topics*, *214*(1), 215–243. doi:10.1140/epjst/e2012-01693-0

Khan, T.M., Clear, F., & Sajadi, S.S. (2012). The relationship between educational performance and online access routines: Analysis of students' access to an online discussion forum. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 226–229). New York, NY: ACM. doi:10.1145/2330601.2330655

Koenker, R. (2008). Censored Quantile Regression Redux. *Journal of Statistical Software*, *27*(6), 1–25.

Koenker, R. (2013). *Quantile Regression in R: A Vignette*. Retrieved from http://www.econ.uiuc.edu/~roger/research/rq/vig.pdf

Mödritscher, F., Andergassen, M., Law, E.L.-C., & García-Barrios, V.M. (2013). Application of learning curves for didactic model evaluation: Case studies. *International Journal of Emerging Technologies in Learning (iJET)*, *8*(S1), 62–69. doi:10.3991/ijet.v8iS1.2357

Mödritscher, F., Neumann, G., & Andergassen, M. (2013). Dependencies between e-learning usage patterns and learning results. In *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies*. Presented at the i-KNOW '13, Graz, Austria.

Rawson, K.A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology. General*, *140*(3), 283–302. doi:10.1037/a0023956

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Buckingham Shum, S., Ferguson, R., … Baker, R.S.J.d. (2011). *Open learning analytics: An integrated & modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques*. SOLAR Society for Learning Analytics Research. Retrieved from http://solaresearch.org/OpenLearningAnalytics.pdf

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*. doi:10.1177/0002764213479366

Srivastava, J., Cooley, R., Deshpande, M., & Tan, P.-N. (2000). Web usage mining: Discovery and applications of usage patterns from Web data. *SIGKDD Exploration Newsletter*, *1*(2), 12–23. doi:10.1145/846183.846188

Thalheimer, W. (2006). *Spacing learning events over time: What the research says*. Retrieved from http://www.work-learning.com/catalog.html

Vlach, H.A., & Sandhofer, C.M. (2012). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development*, *83*(4), 1137–1144.

Wells, R., & Hagman, J.D. (1989). *Training procedures for enhancing reserve component learning, retention, and transfer*. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a217450.pdf

Whitmer, J.C. (2012). *Logging on to improve achievement: Evaluating the relationship between use of the learning management system, student characteristics, and academic achievement in a hybrid large enrollment undergraduate course*. University of California.

Willingham, D.T. (2004). Practice makes perfect — but only if you practice beyond the point of perfection. *American Educator*, *28*(1), 31–33.

Wise, A.F., Zhao, Y., & Hausknecht, S.N. (2013). Learning analytics for online discussions: A pedagogical model for intervention with embedded and extracted analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 48–56). New York, NY: ACM. doi:10.1145/2460296.2460308

Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). Improving retention: Predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 145–149). New York, NY: ACM. doi:10.1145/2460296.2460324