IMPLEMENTATION OF AN ADAPTIVE INSTRUCTIONAL DESIGN FOR A PHYSICS MODULE IN A LEARNING MANAGEMENT SYSTEM

Christof Imhof, Per Bergamin, Ivan Moser and Matthias Holthaus
UNESCO chair on personalized and adaptive distance education, Swiss Distance University of Applied Sciences (FFHS)
Überlandstrasse 12, Brig, Switzerland

ABSTRACT

This article demonstrates how an adaptive instructional design for a physics module can be realized in a standard learning management system. We implemented a didactic design with physics-specific online exercises that were accompanied by either detailed or non-detailed instructions, depending on the results of the previous task (or a prior knowledge test for the very first exercise). This was realized by use of simple technological tools within the framework of a straightforward recommender system with four components. Consequently, students with less prior knowledge and/or lower learning achievements received more and different teaching assistance than those with high levels of prior knowledge or performance. This was done in the form of recommendations embedded within task feedback, suggesting which task to tackle next. We present first results which show that prior knowledge and online activity contribute to the learning progress in different ways depending on the type of task that was chosen. The detailed versions of the tasks were beneficial only to the students with lower or medium prior knowledge test scores while the students with higher levels of prior knowledge had less learning progress. In the future, our simple recommender system may serve as the basis for a more complex adaptive system, further closing the gap between research and practice in the field of technology-based adaptive learning.

KEYWORDS

Technology-based Learning, Adaptive Learning, Cognitive Load, Expertise Reversal Effect, Learning Management System, Log Files

1. INTRODUCTION

As countless aspects of our lives have become more and more digitalized in the past few decades, so has learning. Accordingly, new forms of learning such as distance learning or technology-based learning have emerged and have been gaining in importance ever since (Bergamin et al. 2012). Due to their flexible nature, these new forms of learning allow learners independence and autonomy, offer freedom of choice and break space-time barriers, thus granting many people the opportunity to pursue academic studies in circumstances that do not commonly allow for that to occur (e.g. full-/part-time employment or parenthood). In addition, such flexibility allows for individual requirements and contexts to be taken into account, which can vary substantially between learners and which can then determine the appropriate instructional design. In university education, learners are usually expected to develop the same competences over the course of their studies despite differing in key characteristics such as different prior knowledge or experience in relation to certain topics or learning skills. One way to achieve the goal of comparable learning success despite heterogeneous preconditions is through adaptation of the learning process, replacing the classic "one-size-fits-all" approach.

The importance of adapting the learning process to the needs of the learner can be demonstrated by the finding that instructional techniques (e.g. guidance by a tutor or detailed instructions) that prove beneficial for novices in a given field can lose their effectiveness or even be counterproductive when applied to experts; a phenomenon known as the Expertise Reversal Effect (Kalyuga et al. 2003). On a technical level, adaptive learning environments may be provided through Learning Management Systems (LMS), which are increasingly accessible due to the raise of technology-based learning.

Despite the popularity of studies on adaptive educational hypermedia, actual practical implementations of adaptive technology-based learning are still scarce (Somyürek 2015). Scanlon et al. (2013) found "a surprising failure" (p. 4) to translate research results in the field of technology-based learning (e.g. prototypes) into commercial applications, due to the so-called "valley of death" (i.e. failure due to lack of funding amongst other factors). According to Price et al. (2017), this gap between research and practice in technology-based learning appears to be systemic in nature, requiring change on multiple levels, including institutional change. Murray and Pérez (2015) claim that intelligent technology-based learning environments are still "years away", in spite of the advances that have been made, and appeal for pedagogy rather than technology to "drive the evolution of advanced learning systems" (p. 124). Oxman and Wong (2014) also identify the challenges of (further) implementation of technology-based learning systems as structural (e.g. term length) and operational rather than technological. Therefore, bridging the gap between research and practice requires an interdisciplinary approach that involves large-scale field studies in appropriate contexts as well as well-founded instructional designs (cf. Scanlon et al. 2015). This study seeks to narrow the gap between experimental research and its practical application by addressing the question whether an adaptive learning system can be implemented in a traditional learning environment without the use of high-end technology (such as deep neural networks). In this paper, we demonstrate how adaptive learning can be implemented in practice on a university level by applying it to a physics module within the learning management system "Moodle". Our approach utilises a fairly simple rule-based instructional design. We further explore to what extent adaptive instruction design, online activity and prior knowledge are related and how much they contribute to the learning progress. Finally, we discuss the potential and limitations of rule-based adaptive learning systems within the boundaries of a standard LMS.

2. THEORETICAL BACKGROUND

As explained above, the Expertise Reversal Effect – the phenomenon that teaching support which is beneficial for novices can turn out to be superfluous or even detrimental to experts and vice versa – may obstruct learning success in many learning scenarios, especially in classic settings where all learners receive the same instructions or guidance (Kalyuga 2007b). "Reversal" refers to the circumstance that the effectiveness of didactical aspects may be reversed for different levels of learners' expertise (Lee & Kalyuga 2014). The most common explanation for this effect is the Cognitive Load Theory (Sweller 1988), which focusses on the interactions between long-term and working memory. According to this theory, the former is used to store knowledge and has an unlimited storing capacity while the latter is involved in consciously processing novel information, but is limited in its capacity of storing it, both in terms of amount and duration (van Merriënboer & Sweller 2005). Recent accounts of the theory differentiate between (at least) two kinds of cognitive load, the intrinsic load and the extraneous load. Intrinsic load concerns cognitive processes that are required for processing learning materials and may be affected by the (perceived) complexity or difficulty of the material. In contrast, extraneous load refers to cognitive processes caused by factors that are not directly related to the learning task but are nevertheless crucial for the learning process, for instance convoluted instructional design or unfavourable presentation of the learning material (Kalyuga 2009a).

The basis of the Cognitive Load Theory is the notion that the intrinsic and extraneous loads combined cannot exceed the limitations of the working memory (Paas et al. 2003). If high extrinsic load results from an unnecessary processing of design or presentation aspects, less capacity can be made available for the processing of the actual learning tasks (i.e. intrinsic load). Consequently, learning is impaired if the learning activities require too much cognitive capacity, resulting in overload.

Importantly, the current cognitive load of a learner is not only determined by objective factors (such as difficulty of the learning content and instructional design), but also by characteristics of the learner. In parts, human expertise results from cognitive schemata with varying degrees of complexity and automation housed by the long-term memory (van Merriënboer & Sweller 2005). Knowledge is stored in and organized by these schemata, which may become automated through training, thus freeing space in the working memory, thereby reducing the intrinsic load and leaving more cognitive capacity for the processing of new content (Kalyuga 2009a). Put differently, the level of available knowledge exerts considerable influence on the cognitive load (Kalyuga 2007b). This implies that optimal teaching of complex learning content needs to take learners' cognitive load into account and actively manage it though instructional interventions (Somyürek, 2015; Rey & Buchwald, 2011). In the case of the Expertise Reversal Effect, this concerns the degree of instructional guidance: On the one hand, a lack of sufficient guidance during a complex task may result in the application of poor problem-solving strategies or arbitrary trial-and-error behaviour. On the other hand, vast amounts of instructional guidance may lead learners to squandering their resources by comparing and contrasting their prior knowledge with the incoming information, thus inflating their intrinsic load (Kalyuga 2007b).

Consequently, at the start of the learning process, novices should be provided with instructional guidance (e.g. step-by-step instruction) in order to help them with their tasks and optimise their cognitive load. As the learners gain more expertise over time, this guidance can then gradually be reduced (cf. the concept of fading scaffolds; van Merriënboer & Sluijsmans 2009). The educational implications of the Cognitive Load Theory in general and the Expertise Reversal Effect in particular have been confirmed by numerous studies (e.g. Rey & Buchwald 2011). However, it should also be noted that the cognitive load approach is limited to the acquisition of subject-specific knowledge as the learning goal (Kalyuga & Singh, 2016) and is less applicable to other learning objectives such as the acquisition of self-regulated learning skills or the increase of learning motivation. Moreover, learners may feel overburdened with the necessary monitoring and adaptation of the learning process (Kirschner & van Merriënboer 2013). Technology-based adaptive learning can assist these processes, thus reducing the extrinsic load and increasing the effectiveness of learning.

In contrast to traditional technology-based approaches, adaptive learning allows learning aspects (contents, navigation, support) to be presented in a dynamic environment that continually changes in response to information collected in the course of learning. This raises the question, which sources of a learning scenario are most suitable as a basis for adaptation processes in a course module (cf. Nakić et al. 2015, for an overview). Principally, three main groups of characteristics can be identified: (1) stable or situational personal characteristics of the learners such as gender, culture, learning style, prior knowledge or emotional state, (2) specific characteristics of the content such as topics or task difficulty and (3) characteristics of the context such as learning time or self-regulation (Wauters et al. 2010). The learning process itself can then be adapted by means of altering the instructional design regarding the relevant learning objects in accordance with the needs of an individual or a group of learners.

3. INSTRUCTIONAL DESIGN AND SYSTEM IMPLEMENTATION

In our instructional design, we focus on task difficulty as the adaptive factor, similar to Brunstein et al. (2009); Hsu et al. (2015); and van Der Kleij et al. (2015). Distance students in general and the students at our university in particular tend to have significantly different levels of background knowledge and learning strategies (mostly due to different educational and/or professional careers). Therefore, we implemented an adaptive instruction design in the learning management system used by our university (Moodle), which recommended tasks with step-by-step detailed or non-detailed instructions and thus varied the task difficulty accordingly. In the context of self-regulated learning, students could either follow the recommendation or choose an alternative learning path. Aiming to reduce the cognitive load, less proficient students were given instructions that offered more support and assistance while more proficient students received less support, thus increasing the task difficulty. The recommendations and interventions were each embedded in the feedback of the previously processed task. Our framework for processing and linking learning data with adaptive learning instructions was inspired by a model by Zimmermann et al. (2005). Conceptually, the system was based on four components that together formed the adaptation mechanism. The first component was the sensors, which were linked to the task data (specifically if a task has been solved correctly or not). The second component was the analyser, which collected the data measured by the sensors. This information was then transferred to the third component, the controller, which determined whether a certain threshold had been met. Depending on the outcome, the controller determined if the learning object (for example a task) was to be adjusted. The last component was the presenter, which then displayed objects of learning support (such as recommendations). As an entry point, we used the data from a prior knowledge test that was administered at the beginning of the course and assessed the level of expertise with which the students started the course. In the course of the semester, further tests and assessments were then fed to the sensor component data base. Consequently, students received instructions and learning support adapted to their learning performance and behaviour. The learning support focused on three different elements: the initial sensor, the step loop and the task loop.

The initial sensor was an assessment of prior knowledge in physics in the form of a set of standard exercises solved at the very start of the course. Based on their performance, the students were then divided into two groups, “novices” (less than 50% correct answers) and “experts” (more than 50% of the answers were correct). Depending on the score, the first proper task in the module appeared in a detailed (high learning support) or a standard form (low learning support). The second element, the step loop, measured the current level of knowledge within a task and accordingly determined the appropriate learning support. Based on the correctness of the answer, the students received feedback which was provided after each step in the task and changed depending on how often the same question was answered incorrectly. This served the purpose of clarifying possible misunderstandings the students might have had as quickly as possible, e.g. by reminding them of forgotten information (Durlach & Ray 2011). The third element, the task loop, consisted of two kinds of tasks, namely standard tasks and transfer tasks. The former assessed the ability to solve a particular physics problem, while the latter had two objectives: On the one hand, the transfer task checked whether a student had understood a particular problem and was thus able to solve the task in its standard form (see “vertical transfer”, van Eck & Dempsey 2002), and on the other hand, it evaluated whether the student was able to apply the now acquired problem-solving knowledge to a similar task in a different topic (see “horizontal transfer”, van Eck & Dempsey 2002). Within each set of tasks, the system recommended which task the student should tackle next and in what form (detailed or standard). The detailed version featured numerous small solution steps, while the standard version was composed of few solving steps. We chose a rule-based adaptive learning system with a fixed set of rules. The reason for this decision was two-fold: On the one hand, the sensors in our learning scenario do not generate enough data for a complex self-learning system and on the other hand, we wanted to keep the adaptation mechanisms transparent for our students in the sense of an “open learner model”.

Starting in the autumn semester of 2015/16, we carried out a two-year field study that implemented our adaptive learning approach into one of our university’s study modules. The chosen bachelor module was part of the course of studies in industrial engineering and featured three main physics topics (thermodynamics, optics and microphysics). The module was organized in a blended learning format, which means there was a mixture of face-to-face sessions (20% of the overall expected effort) and both on- and off-line self-study (80%). In order to promote the acceptance of our system among students, we chose a mixed control approach between the system (adaptation) and the learner (adaptability). In their feedback, the students thus received recommendations as to which tasks they should ideally complete and in which form they should choose it. As previously stated, compliance with these recommendations was always optional. The sensor data (i.e. the current state of knowledge) was made available to students in a transparent and concise way to foster their self-assessment skills as well as their acceptance of the recommendations (see open learning models, e.g. Long & Aleven 2017; Suleman et al. 2016).

4. ANALYSIS

4.1 Object of Investigation, Subjects, Procedure and Hypotheses

As previously stated, the students attending a physics module in the semesters 2015/16 and 2016/17 served as the participants in our investigation. The module is offered each autumn semester and was chosen for its reputation as a “problem module”, due to its above-average failure rate. Each semester, the students are divided into seven or eight classes, split among different lecturers. Considering the high degrees of employment of our students, the class division was based on the students’ preferences in terms of optimal time and location for the five face-to-face events (with four options to choose from in terms of location), which had no bearing on the module content or the online part of the course. In the autumn semester 2015/16, 105 students were enlisted in the course while 106 students participated in the course the following year (2016/17). In both years, the data of several students had to be excluded for our main analyses since they either didn’t complete the prior knowledge test (11 in the 2015/16 semester) or the final exam at the end of the semester (7 in the first and 16 in the second year). There were no missing prior knowledge test scores in the second year since the test was not optional anymore. The course offered 43 adaptive tasks not all of which had to be completed by the students. More proficient students for example may only need to solve a fraction of the task array after having successfully completed the first assessment (initial sensor) in order to succeed at the final test at the end of the semester while less proficient students may have to complete a larger portion of the task set in order to achieve the same goal (following the recommendations).

As for the procedure, we first evaluated the distributions of the prior knowledge test scores and three different indicators of online activity (sum of daily clicks, sum of completed standard tasks and sum of completed detailed tasks) based on the students' log files. The sum of daily clicks was used as a general online activity measure, while the other two were specific for engagement with either detailed or non-detailed tasks. In order to investigate how our instruction design impacted the learning progress of our participants, we then explored the relationship between online activity, prior knowledge and the learning progress by calculating several regression models. We formulated four hypotheses for this exploratory study: H_1: prior knowledge is negatively related with learning progress (the less one knows, the more one can learn, i.e. a ceiling effect); H_2: general online activity (in the form of the sum of daily clicks) and learning progress are positively related; H_3: engagement with the tasks (be it detailed or standard tasks) is positively related with learning progress;

H_4: There is a negative interaction between online activity (all three forms thereof) and prior knowledge (more online activity benefits less knowledgeable students more than it does those with high levels of prior knowledge). The whole procedure was done separately for both semesters since the second semester served as a replication and expansion of the first. All statistical analyses were performed with R (R Core Team 2013).

4.2 Distributions of the Prior Knowledge Test Scores and Online Activity Measures

Before examining the distribution of the prior knowledge test scores, the scores in the separate three topics (microphysics, thermodynamics and optics) were added and standardised to an index of 100. Using the R psych package (Revelle 2017), we calculated the mean scores of the prior knowledge test and the mean sums for the online activity measures, as well as the skewness and kurtosis of the distributions (see *Table 1* for an overview). The normality of the distributions was then tested using the Shapiro-Wilk normality test. In the 2015/16 semester, the null hypothesis that the data was normally distributed wasn't rejected (W = 0.99, p = .59) while it clearly was in the 2016/17 semester (W = 0.94, p < .001). Thus, for the semester of 2015/16, the test scores were found to be approximately normally distributed while in the 2016/17 semester, the data was not normally distributed but instead skewed to the left due to overall low test scores (not a single person reached a score higher than 45 out of 100) and a high proportion of test scores of 0.

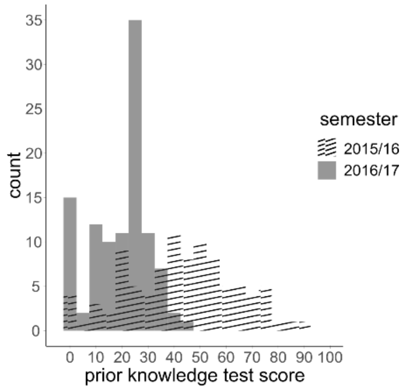

Figure 1. Two overlapping histograms showing the distributions of the prior knowledge test scores in both semesters. The frequencies of the scores did not differ from a normal distribution in year 1 (2015/16) but were left-skewed in year 2 (2016/17)

ISBN: 978-989-8533-81-4 © 2018

As for the online activity measures, only completed tasks were accounted for since the recommendation in the task loop was only given after having completed a task; aborted tasks were not counted. All six online measure distributions were clearly left-skewed due to the high frequencies of low levels of activity.

Table 1. Distribution of the prior knowledge test scores and three online activity measures in the two semesters, including the number of participants, means, standard deviations, the Shapiro-Wilk test statistic W and the according p-value

Measure	Mean (*SD*)	Skewness	Kurtosis	W (*p*)
	Semester 2015/16			
Prior knowledge test score (n=94)	42.88.0 (20.93)	-0.06	-0.53	0.99 (.59)
Sum of daily clicks (n=87)	831.56 (693.19)	1.25	1.77	0.89 (<.001)
Sum of completed standard tasks (n=87)	7.03 (8.72)	1.40	1.26	0.80 (<.001)
Sum of completed detailed tasks (n=87)	9.25 (8.30)	0.65	-0.72	0.90 (<.001)
	Semester 2016/17			
Prior knowledge test score (n=106)	19.71 (11.02)	-0.36	-0.67	0.94 (<.001)
Sum of daily clicks (n=90)	860.75 (626.45)	0.61	-0.07	0.95 (<.001)
Sum of completed standard tasks (n=90)	11.10 (9.55)	0.75	0.11	0.92 (<.001)
Sum of completed detailed tasks (n=90)	7.85 (7.32)	0.76	-0.10	0.90 (<.001)

4.3 Prior Knowledge, Online Activity and Learning Progress

In the following section, we report the relationship between prior knowledge, online activity and the learning progress within the adaptive module. For this purpose, the learning progress for each student was defined as the difference between the results of the prior knowledge test and the final test, both standardized to 100. We explored the relationship between the prior knowledge, the students' online activity and their learning progress by calculating three regression models for each of the three online activity measures ("OA" models 1 to 3) as well as a regression model solely containing prior knowledge (the "PK" model), all of which was done separately for both semesters. We used the R package ggplot2 (Wickham 2009) for the regression plots. The first OA model included only the online activity measure as a predictor. The second OA model added a second predictor in the form of the prior knowledge test score and the third and final OA model added the interaction between the online activity measure and the prior knowledge test score (see *Figure 2*). In all instances, the three models were supposed to predict the learning progress.

4.3.1 Prior Knowledge

For both semesters, the PK model yielded a significant regression with an adjusted R^2 of 0.29 for the 2015/16 semester ($F(1, 85) = 36.01$, $p < .001$) and an adjusted R^2 of 0.03 for the 2016/17 semester ($F(1,85) = 4.02$, $p = .047$). In line with our hypothesis H_1, there was a significant negative relationship between prior knowledge test score and predicted learning progress ($b = -0.61$, $p < .001$) for the 2015/16 semester, i.e. students with higher prior knowledge showed less learning progress. A similar relationship was found for the 2016/17 semester ($b = -0.31$, $p = .048$).

4.3.2 Sum of Daily Clicks and Prior Knowledge

For the 2015/16 data, two significant regression models were found, namely OA models 2 ($F(2, 84) = 21.8$, $p < .001$) with an adjusted R^2 of 0.33 and model 3 ($F(3, 83) = 15.82$, $p < .001$) with an adjusted R^2 of 0.34. In model 2, there was a positive relationship between sums of daily clicks and predicted learning progress ($b = 0.01$, $p = .02$) and a negative relation between prior knowledge and learning progress ($b = -0.67$, $p < .001$), supporting H_2 and H_1 respectively. Contrary to H_4, the interaction between the two predictors (model 3, see *Figure 2*) was not significant ($b = -0.0003$, $p = .09$). Model 2 significantly improved model fit compared to model 1 ($F(1, 84) = 43.96$, $p < .001$), while model 3 did not did not improve model fit compared to model 2 ($F(1, 83) = 2.88$, $p = .09$). In contrast, none of the regression models were significant for the 2016/17 data.

4.3.3 Sum of Completed Standard Tasks and Prior Knowledge

We again found two significant regression models in the 2015/16 semester, namely for OA models 2 ($F(2, 84) = 29.07$, $p < .001$) and 3 ($F(3, 83) = 19.18$, $p < .001$), both with an R^2 of 0.39. In model 2, both factors were significant predictors of learning progress. As before, participants' predicted learning progress was positively related to the sum of completed standard tasks (b = 1.08, $p < .001$) and negatively related to the prior knowledge test score (b = -0.91, $p < .001$). In model 3 (see *Figure 2*), the interaction between the two predictors again did not reach significance (b = -0.003, $p = .80$). Thus, the results supported hypotheses H_1 and H_3 but not H_4. As before, model 2 provided a better model fit ($F(1, 84) = 56.51$, $p < .001$) compared to model 1. The same could not be said about model 3 ($F(1, 83) = 0.06$, $p = .80$). Again, none of the regression equations were significant in case of the 2016/17 data.

4.3.4 Sum of Completed Detailed Tasks and Prior Knowledge

In the 2015/16 semester, all three models were significant, with OA model 1 ($F(1, 85) = 6.54$, $p = .01$) having an R^2 of 0.06, model 2 ($F(2, 84) = 21.2$, $p < .001$) an R^2 of 0.32 and model 3 ($F(3, 83) = 17.67$, $p < .001$) an R^2 of 0.37. In model 1, the predicted learning progress was positively related to the sum of completed detailed tasks (b = 0.75, $p = .01$). This was also in the case in model 2 (b = 0.55, $p = .03$). Like before, the relationship between predicted learning progress and prior knowledge test was negative (b = -0.58, $p < .001$). Unlike before, the interaction between online activity and prior knowledge (see *Figure 2*) was significant (b = -0.04, $p = .01$). Hence, hypotheses H_1, H_3 and H_4 were all supported by the data. A simple slope analysis (using the R package jtools, Long 2018) revealed that the slope of the sum of completed detailed tasks was only significantly different from zero when the prior knowledge test score was either medium (b = 0.52, $p = .04$) or low (b = 1.28, $p < .001$). As before, model 2 ($F(1, 84) = 35.92$, $p < .001$) was an improvement over model 1 in terms of model fit. Unlike before, model 3 further improved upon model 2 in this case ($F(1, 83) = 7.39$, $p =.01$). However, none of the regression models were significant for the 2016/17 data.

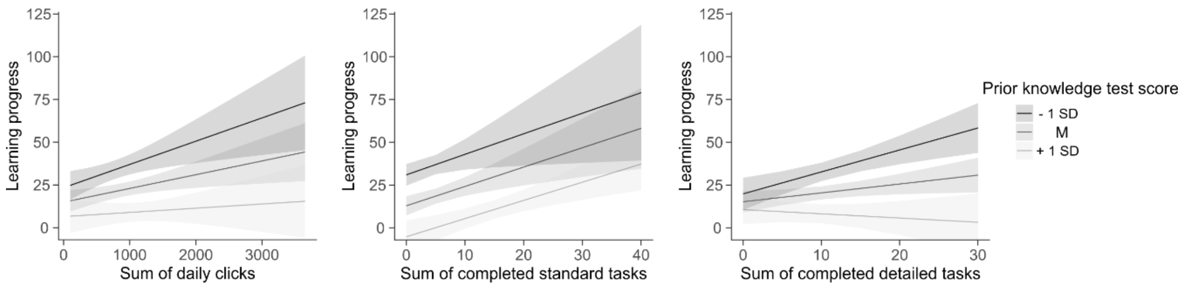

Figure 2. Regression model 3 for all three measures of online activity separately. All models are based on 2015/2016 data exclusively. In model 3, learning progress is predicted by *prior knowledge test score*, *online activity* and the interaction between the two. That interaction was significant only when using sum of completed detailed task as the measure of online activity. In order to illustrate the interaction, we plotted the predicted effect of online activity on learning progress using simple slopes for students that scored one standard deviation below the mean (-1 SD), students that scored one SD above the mean (+1 SD) and students with mean prior knowledge test scores (M)

5. DISCUSSION

The purpose of this article was to demonstrate how an adaptive instruction design could be implemented in a standard learning management system. On a theoretical level, the instruction design was based on the Cognitive Load Theory in general and the Expertise Reversal Effect in particular. We developed an adaptive task set which was then combined with a rule-based recommendation system within the learning management system Moodle. Our adaptive system was applied to a physics module during two semesters (2015/16 and 2016/17) with the respective enlisted students of our university serving as our participants.

The results for the 2015/16 semester reveal that the prior knowledge test score and the level of online activity are both important predictors of learning progress. As hypothesized, the relationship between online activity (be it the sum of daily clicks or the amount of standard tasks solved) and learning progress was positive (the higher the level of online activity, the larger the learning progress), while the relationship between prior knowledge test scores and learning gain was negative (the higher the initial score, the smaller the learning gain). Contrary to our hypotheses, this effect was independent from the level of online activity in most cases with the detailed tasks being the exception, where we found the expected negative interaction: the amount of tasks solved was positively related with learning gains when the prior knowledge test score was low and negatively related when the score was high. This means the less proficient students (i.e. the "novices") benefitted from the detailed tasks more than the advanced students (i.e. the "experts"). Even though this result suggests a presence of the Expertise Reversal Effect (Kalyuga et al. 2003), an according conclusion cannot be drawn unless we know whether the students followed the recommendations. Moreover, the ceiling effect we found could also account for that result, especially when considering that classic exams always have a maximum score that cannot be exceeded.

Surprisingly, neither the prior knowledge test scores nor the levels of online activity predicted the learning progress in the 2016/17 semester, neither separately nor in interaction. We argue that this result was found due to the peculiar distribution of the prior knowledge test scores. As stated before, the scores in that semester were very low with not even a single student reaching 50 points out of a 100. The test was exactly the same as the year before, where it showed no floor or ceiling effects and yielded a mean slightly below what one would expect for a test of medium difficulty (in this case, 50). Since the test itself was identical, the difference between the semesters must have a different explanation. Even though the distributions of the online activity levels were also left-skewed, this could not explain the difference since those distributions were similar in both semesters. As previously stated, what differed between the semesters was the circumstance that the prior knowledge test was mandatory in the 2016/17 while it was not the year before. Motivational aspects thus may factor into a possible explanation or a potential homogeneity within the particular group of students, which has yet to be investigated (e.g. similar previous education that was light on physics).

A major limitation of the study is the lack of a control group. Control groups are hard to achieve in contexts like the one this field study was conducted in since the grades the students receive at the end of the semester are real, thus posing ethical problems comparable to withholding of treatment in clinical studies. In this particular case, there was no non-adaptive past version of the course available to compare to the adaptive one, and even if there was, comparisons might not be all that conclusive given the difference in key factors we witnessed between the 2015/16 and 2016/17 semesters. As of the time of this writing, the adaptive system is still being used in the 2017/18 semester, the results of which will be available soon. Another limitation is the lack of information concerning whether the students actually complied with the recommendations (rather than making their choice either at random or as they saw fit). However, the students' compliance with the recommendations will be analysed soon as well.

6. CONCLUSION

We demonstrated that a simple rule-based adaptive system can be implemented in a common learning management system, getting one step closer to bridging the gap between theory and practice of implementing adaptive systems. As our results show, the implementation of the task-difficulty based adaption process was successful (to an extent), even though we (expectedly) encountered a ceiling effect concerning the learning progress of the students with higher levels of prior knowledge. Our results also demonstrate the importance of reliable assessments serving as the sources of adaption (e.g. well-balanced tests). Future studies in this line of research could grant insight into the transfer tasks, which were very rarely solved by the students and thus could not be analysed properly. Despite the limitations of this study, this line of research holds potential since the successful implementation of simple rules can serve as the basis for a more complex adaptive system, for instance by expanding the sources for adaption (e.g. mood) or by refining the instruction design (e.g. by adding more levels of difficulty). In the near future, research in the field of learning analytics (see e.g. Bannert et al. 2017) may result in the development of more holistic sensors that measure several metrics relevant to learning at once, providing multiple sources for adaption, potentially even allowing implementations within common LMS. Such improvements could result in adaptive systems with more intelligent and complex algorithms at their core, which could then lead to more accurate prognoses, higher efficiency of the adaptive process and more suitable recommendations for the students, which in turn may enhance the acceptance of such systems among students and lecturers.

ACKNOWLEDGEMENT

We would like to thank the many students who participated in the course as well as Dr. Stéphanie McGarrity for proofreading this manuscript and giving us invaluable advice.

REFERENCES

Bannert, M. et al., 2017. Relevance of Learning Analytics to Measure and Support Students' Learning in Adaptive Educational Technologies. In *LAK'17 Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. Vancouver, BC, Canada: ACM New York, NY, USA, pp. 568–569.

Bergamin, P.B. et al., 2012. The Relationship between Flexible and Self-Regulated Learning in Open and Distance Universities. *International Review of Research in Open and Distance Learning*, 13(2), pp.101–123.

Brunstein, A., Betts, S. & Anderson, J.R., 2009. Practice Enables Successful Learning Under Minimal Guidance. *Journal of Educational Psychology*, 101(4), pp.790–802.

Durlach, P.J. & Ray, J.M., 2011. *Designing adaptive instructional environments: Insights from empirical evidence (Technical Report 1297)*, Arlington, VA: U. S. Army Research Institute for the Behavioral Social Sciences.

van Eck, R. & Dempsey, J., 2002. The Effect of Competition and Contextualized Advisement on the Transfer of Mathematics Skills in a Computer-Based Instructional Simulation Game. *Educational Technology Research & Development*, 50(3), pp.23–41.

Hsu, Y. et al., 2015. Interactions Between Levels of Instructional Detail and Expertise When Learning with Computer Simulations. *Journal of Educational Technology & Society*, 18, pp.113–127.

Kalyuga, S., 2007b. Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction. *Educational Psychology Review*, 19(4), pp.509–539.

Kalyuga, S., 2009. Managing cognitive load in adaptive ICT-based learning. *Journal of Systemics, Cybernetics and Informatics*, 7(5), pp.16–21.

Kalyuga, S. et al., 2003. The Expertise Reversal Effect. *Educational Psychologist*, 38(1), pp.23–31.

Kalyuga, S. & Singh, A.-M., 2016. Rethinking the Boundaries of Cognitive Load Theory in Complex Learning. *Educational Psychology Review*, 28(4), pp.831–852.

van der Kleij, F.M., Feskens, R.C.W. & Eggen, T.J.H.M., 2015. Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes: A Meta-Analysis. *Review of Educational Research*, 85(4), pp.475–511.

Lee, C.H. & Kalyuga, S., 2014. Expertise reversal effect and its instructional implications. In *Applying science of learning in education: Infusing psychological science into the curriculum*. Washington, DC: Society for the Teaching of Psychology, pp. 32–44.

Long, J.A., 2018. *jtools: Analysis and Presentation of Social Scientific Data*, R package version 1.0.0, https://cran.r-project.org/package=jtools.

Long, Y. & Aleven, V., 2017. Enhancing learning outcomes through self-regulated learning support with an Open Learner Model. *User Modeling and User-Adapted Interaction*, 27(1), pp.55–88.

van Merriënboer, J.J.G. & Sluijsmans, D.M.A., 2009. Toward a Synthesis of Cognitive Load Theory, Four-Component Instructional Design, and Self-Directed Learning. *Educational Psychology Review*, 21(1), pp.55–66.

van Merriënboer, J.J.G. & Sweller, J., 2005. Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, 17(2), pp.147–177.

Murray, M.C. & Pérez, J., 2015. Informing and performing: A study comparing adaptive learning to traditional learning. *Informing Science*, 18(1), pp.111–125.

Nakić, J., Granić, A. & Glavinić, V., 2015. Anatomy of student models in adaptive learning systems: A systematic literature review of individual differences from 2001 to 2013. *Journal of Educational Computing Research*, 51(4), pp.459–489.

Oxman, S. & Wong, W., 2014. White Paper: Adaptive Learning Systems. *DV X Innovations DeVry Education Group*.

Paas, F., Renkl, A. & Sweller, J., 2003. Cognitive Load Theory and Instructional Design: Recent Developments. *EDUCATIONAL PSYCHOLOGIST*, 38(1), pp.1–4.

Price, L., Casanova, D. & Orwell, S., 2017. Modeling an institutional approach to developing Technology Enabled Learning: Closing the gap between research and practice. *INTED2017 Proceedings*, pp.5009–5018.

R Core Team, 2013. R: a language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Revelle, W.R., 2017. psych: Procedures for Personality and Psychological Research.

Rey, G.D. & Buchwald, F., 2011. The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, 17(1), pp.33–48.

Scanlon, E. et al., 2013. *Beyond prototypes: Enabling innovation in technology-enhanced learning*, Open University, Milton Keynes.

Scanlon, E., O'Shea, T. & McAndrew, P., 2015. Technology-Enhanced Learning : Evidence-based Improvement. In *L@S '15 Proceedings of the Second (2015) ACM Conference on Learning @ Scale*. Vancouver, BC, Canada: ACM New York, NY, USA, pp. 229–232.

Somyürek, S., 2015. The new trends in adaptive educational hypermedia systems. *International Review of Research in Open and Distributed Learning*, 16(1), pp.221–241.

Suleman, R.M., Mizoguchi, R. & Ikeda, M., 2016. A New Perspective of Negotiation-Based Dialog to Enhance Metacognitive Skills in the Context of Open Learner Models. *International Journal of Artificial Intelligence in Education*, 26(4), pp.1069–1115.

Sweller, J., 1988. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12, pp.257–285.

Wauters, K., Desmet, P. & Van den Noortgate, W., 2010. Adaptive item-based learning environments based on the item response theory: possibilities and challenges. *Journal of Computer Assisted Learning*, 26, pp.549–562.

Wickham, H., 2009. *Ggplot2 : elegant graphics for data analysis*, Springer.