Towards motivation-based adaptation of difficulty in e-learning programs

Anke Endler
Julius-Maximilians-Universität

Günter Daniel Rey
FernUniversität in Hagen

Martin V. Butz
University of Tübingen

The objective of this study was to investigate if an e-learning environment may use measurements of the user's current motivation to adapt the level of task difficulty for more effective learning. In the reported study, motivation-based adaptation was applied randomly to collect a wide range of data for different adaptations in a variety of motivational states. This data was then utilised to extract rules for an adequate motivation-based adaptation to maximise expected learning success. A learning classifier system was used for the data analysis, generating rules for suitable and unsuitable adaptations based on current user motivation data. We extracted a set of twelve rules which suggest particular adaptation strategies based on real-world data. These rules could generally be embedded into existing psychological theories, namely the Zone of Proximal Development and the Yerkes-Dodson Law. In future research, we intend to evaluate these rules on further studies and develop concrete sets of adaptation strategies based on user motivation measurements.

Introduction

E-learning is a large part of the current educational process. A growing number of educational institutions utilise e-learning, such as learning games suitable to teach a particular (small) topic, online language courses, and even entire university programs using virtual classrooms. Some of these e-learning environments, free and commercial, are used in real instructional settings. Self-study language programs are a well-known example. However, most of these instructional systems still employ a 'one-size-fits-all' philosophy, where every learner, irrespective of their individual abilities and needs, is presented with the same content in the same way (e.g., Rey, 2009). Although these learning environments may help students to learn, adapting e-learning environments to the individual needs of each student may help learners reach their learning goal even more quickly and effectively (e.g., Chen, Lee & Chen, 2005; Shute & Towle, 2003).

These adaptive e-learning environments have been researched for quite a while. According to Mödritscher, Garcia-Barrios and Gütl (2004), the macro-adaptive approach, which allows the adaptation of some main components, can be traced back to the 1970s. Current developments in the field of computer science make it possible to

build and thoroughly test ever more complex learning environments, which may take even more adaptation criteria into account. One of these criteria is the learner's current motivation.

Motivation is an internal condition that activates, energises, and directs behaviour. It has been shown to correlate with learning performance in several studies. Schiefele, Krapp and Winteler (1992) concluded from a meta-analysis that measured interest can predict school achievement. Using a program that teaches tenth-grade students genetics, Song and Keller (2001) showed that a motivationally adaptive strategy promised a significantly higher performance and efficiency, compared to a motivationally minimised or saturated strategy. Keeley, Zayac and Correia (2008) conducted a study that showed a statistically significant, nonlinear relationship between anxiety and test performance of students enrolled in an introductory statistics course.

We use a questionnaire to assess motivation in our study (cf. Rey & Buchwald, 2011), and a learning classifier system (LCS) to evaluate the influence of motivation on learning in the scope of a simple e-learning environment. LCSs are adaptive systems (Holland, 1975) which can learn general rules from exemplary input, using reinforcement learning (Sutton & Barto, 1998) and genetic algorithms (Goldberg, 1989). Given enough input values and an extractable underlying data structure, LCSs are able to learn rules that are expected to maximise predicted payoff (for more information on LCSs, see e.g. Butz, 2006). Moreover, the extracted rules can be used as a basis for developing fundamental motivation-based adaptation strategies in e-learning systems.

The results of this work may help us understand the influence of a learner's current motivation on their learning success. Consequently, better e-learning systems may be built, which take these insights into account to better adapt to a user's motivation, and thus to help them learn more quickly and efficiently.

We first discuss different aspects of motivation, motivational design and motivation-based adaptation. Next, we consider different theories relevant for learning and motivation, and introduce the conducted study in detail and present the generated results. We then introduce the LCS that was used to extract adaptation rules from the available data and look at the extracted rules in more detail, embedding them into well-known theories like the *Zone of Proximal Development* (ZPD) and the *Yerkes-Dodson Law*. Current limitations and future directions conclude this work.

Motivation

Motivation-based adaptation (cf. Keller, 2006), as understood in this paper, aims to improve learning success by adapting the learning environment or learning tasks based on the learner's current motivation. It does not aim primarily to increase the learner's motivation, although it may try to influence the learner's motivation if motivation in turn promises to influence learning success. In short, motivational design aims to influence motivation, while motivation-based adaptation aims to utilise knowledge on motivation to influence learning success.

Another important aspect of the integration of motivational factors into adaptive e-learning is motivation detection, which concerns itself with the assessment of

motivation. Del Soldato and du Boulay (1995) and de Vicente (2003) discuss several approaches to measure motivation in e-learning environments. Questionnaires at the beginning of the interaction may gather relevant information such as the learner's confidence state, but the questionnaire's static nature prevents the consideration of motivational changes during the use of the program. De Vicente's (2003) approach of self-report during the interaction is similar to questionnaires, but more dynamic. Students can manipulate sliders to indicate their motivational state at any time during the use of the program. However, students may forget to update the sliders during their interaction with the program, making this approach potentially unreliable. Another dynamic approach is the utilisation of standard reactions such as a request for help, or the act of giving up. Paralinguistic cues, such as facial expressions, body language, or intonation, can also indicate a learner's motivation and used for adaption (Sarrafzadeh, Alexander, Dadgostar, Fan & Bigdeli 2008). However, the assessment of these cues might be laborious and may be considered intrusive by the learner.

ZPD, flow and the Yerkes-Dodson law

Two well-known theories which predict the influence of learner's expertise and leaner's arousal on learning success are the ZPD (Vygotski, 1963) and the Yerkes-Dodson Law (Broadhurst, 1957; Yerkes & Dodson, 1908). These theories might be adaptable to the influence of motivation on learning success. The original concept of the ZPD predicts that the highest learning success can be expected if expertise and difficulty are on a similar level (cf. Figure 1). If difficulty is too high or too low for the learner's current expertise, learning success will decrease. The theory of flow (Csikszentmihalyi, 1990) assumes that high difficulty for low expertise will lead to anxiety, while low difficulty for high expertise will lead to boredom. The assumptions from the concept of the ZPD and the theory of flow are integrated in Figure 1. The figure includes anxiety, boredom and ZPD as a function of task difficulty (y-axis) and learner's expertise (x-axis).
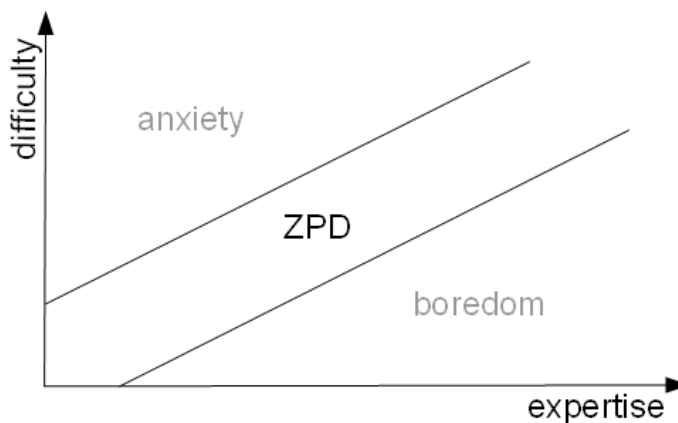

Figure 1: An integrated model including the concept of the zone of proximal development (ZPD) and theory of flow (Csikszentmihalyi, 1990; Vygotski, 1963).

The Yerkes-Dodson Law assumes an interrelation between arousal, task difficulty and performance. The law postulates that a certain amount of arousal is necessary to

activate learning. If arousal is low, performance will also tend to be low. The optimal arousal, i.e. the level of arousal that will lead to maximal performance, is on an intermediate level across the possible range. Higher arousal will lead to a decrease of performance again, so that the correlation between arousal and performance forms an inverted U-shape. For easy tasks, higher activating arousal is expected to result in higher performance compared to difficult tasks, where the optimum is reached for a medium level of arousal. Assuming that arousal and motivation overlap (remember that motivation is defined as internal condition that activates, energises, and directs behaviour), the Yerkes-Dodson Law can be applied to research considering the influence of motivation on learning success.

The present study

Previous research suggests that motivation plays an important role in learning (Keeley et al. 2008; Song & Keller, 2001) and e-learning (e.g., Kim & Frick, 2011). The objective of our study was to find plausible strategies for the adaptation of task difficulty, given learners' current motivations. We aimed to find the adaptation strategies for an adaptive e-learning environment which promise the best expected learning success.

Method

Participants

In total, 37 participants took part in the study, which took place in August 2010. Of the participants, 38% (14) were male and 62% (23) female. Ages ranged from 16 to 48 years, with an average age of 26.7 years (SD = 5.8 years). Twenty-one of the participants were students, four were pupils, three were working on their doctorates, two were currently unemployed, one was in early retirement, and six were working in various areas of expertise. Participants each received 6 Euros for their study participation.

Program, material, instruments, and tasks

ELMA (E-Learning with Motivational Adaptation) is a micro-adaptive e-learning program, which has been developed for this study. It presents a fixed number of tasks and measures the learner's motivational state several times during learning. A simple adaptation, which reacts to the learner's current performance, tries to keep the task difficulty on a level that is consistent with the learner's expertise to avoid the negative impact of the expertise reversal effect (Kalyuga, Ayres, Chandler & Sweller, 2003; Kalyuga, 2007). For the present study, a second adaptation was done randomly after each assessment of the learner's motivational state in order to collect a broad sample of data. This data was then used to extract rules for an adequate motivation-based adaptation.

ELMA worked with a very simple interface, which consisted mainly of black text on white background (cf. Figure 3). Some important information was highlighted using larger or bold letters. An animated progress bar indicated the time remaining for the learner to solve a task. This progress bar also used colors (green and red) to signal whether time was running short.

The flow of the program can be seen in Figure 2. In the following sections, the various parts of the program are presented in more detail.
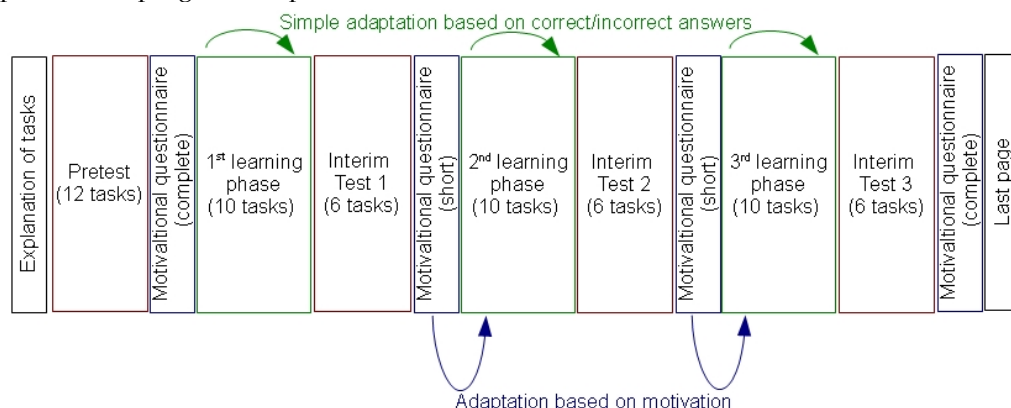

Figure 2: The flow of the e-learning program ELMA

Text pages were used at the beginning and also at the end of the program to round off the experiment. The text pages gave the learner information about the flow of the program, the tasks, and about what we hoped to gain by their participation in the study.

We used self-assessed motivation via the "Questionnaire on Current Motivation" (QCM) which was introduced and evaluated by Rheinberg, Vollmeyer and Burns (2001) in our study. Satisfactory reliabilities for the QCM have been demonstrated for both the German and American versions with six samples ($N = 944$). Furthermore, several studies have supported the questionnaire's validity by showing that initial motivational factors were related to subsequent learning behaviours and outcomes (Rheinberg et al., 2001). The questionnaire consisted of 18 questions and four subdimensions (anxiety, probability of success, interest, and challenge), all containing 7-point Likert scales. In the questionnaire, the learner was asked to report their current motivation based on the previous block of tasks.

The complete questionnaire assessed the learner's motivation at the beginning and at the end of the program, but was not considered for the motivation-based adaptation. For the second complete questionnaire the statements were altered slightly, to make them suitable for a retrospective view on the tasks.

A reduced version of the motivational questionnaire was presented several times during the program to make sure that the program always captured the learner's current motivation. This reduced questionnaire consisted of four statements, covering each of the motivational factors. We chose the statements that had proven to give the best indication for each factor in the study conducted by Rheinberg et al. (2001). In a principal component analysis, each of these statements showed a correlation of at least 0.75 with the respective motivational factor (.85 for *anxiety*, -.78 for *probability of success*, .79 for *challenge*, and .75 for *interest*). These statements were "Es ist mir etwas peinlich hier zu versagen." ("I'm a little embarrassed to fail here.", .85) for *anxiety*; "Wahrscheinlich werde ich die Aufgabe nicht schaffen." ("I probably won't be able to

solve this task.", -.78) for *probability of success*; "Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde." ("I'm very curious as to how well I will perform here.", .79) for *challenge*; and "Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten." ("I would work on a task like this in my free time.", .75) for *interest*.

The reduced motivational questionnaire followed after an interim test and was succeeded by a learning phase. The next interim test evaluated the quality of the chosen adaptation, using the interim test before the questionnaire to exclude the learner's prior knowledge., i.e. the learning success was defined by the percentage of correct answers in the current test compared to the percentage of correct answers in the test before the last learning phase. The questionnaire was presented immediately before the corresponding learning phase to measure the student's current motivation without any delay caused by instructions or tasks that may influence the learner's motivational state.
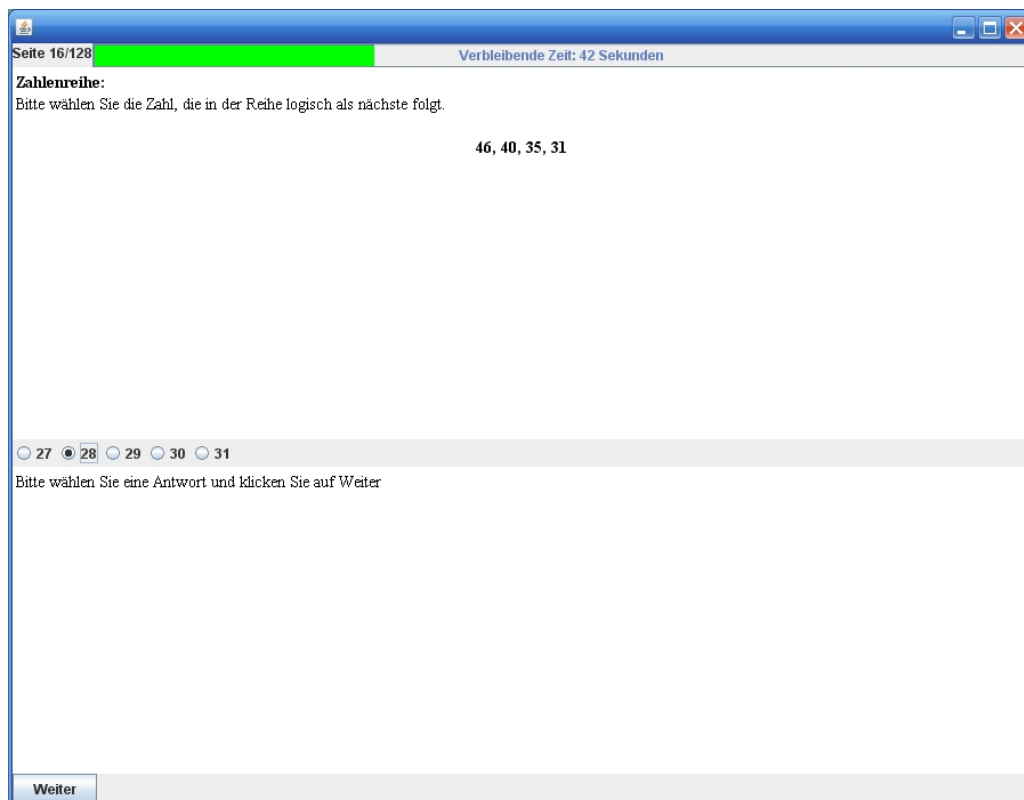

Figure 3: The screenshot shows a task (type 0 task, i.e. the sequence of numbers task) of ELMA, which was presented to the participants

Every task consisted of a presentation, five response options, and feedback, which explained the correct answer. When presented with a task, the presentation and the five response options were shown to the learner. The learner then had to choose the right answer out of these five options. Accordingly, there was a 20% chance that a learner would choose the right answer if they chose one randomly. No aids such as a

calculator or pen and paper were allowed when solving the tasks. A time limit for each task was set to 60 seconds. A green progress bar at the top of the program window reminded the learner of how much time they had left to solve the task. The progress bar turned red when time ran low, i.e., when less than 25% of the time was left.

Once the learner chose an answer and clicked the continue button, or time ran out, the learner received feedback for this specific task. This feedback consisted of whether the task had been solved correctly by the learner, the correct answer, and an explanation of the correct answer.

The program presented two types of tasks in six levels of difficulty. The types of tasks, and the tasks themselves, were taken from the *Kognitiver Fähigkeitstest* (cognitive capability test, KFT, Heller & Perleth, 2000). The test was originally constructed as a cognitive capability test for secondary school students, but, as the average percentage of correctly solved tasks in our study confirms, their demand on mathematical abilities proved challenging for adults as well. We used the tasks Q2, called type 0 here, and Q3, which we call type 1. In the study, the types of task were not adapted, but simply alternated to give the learner some variety and avoid boredom.

In the type 0 task, also known as the sequence of numbers task, the learner was presented with numbers, and occasionally letters. The sequence was determined by a certain rule. The learner had to detect the (most obvious) rule and choose the next number of the sequence from the presented answers.

The type 2 task, also known as the equation task, presented the learner with several numbers and mathematical operators. They had to combine these numbers and operators to form an equation, the result of which was one of the presented answers. All the numbers and operators presented had to be part of the equation. Operators ranged from addition, subtraction, multiplication, and division to bracket terms, square roots, and the power of two in the higher levels of difficulty.

The pretest consisted of 12 tasks, one of every type and level of difficulty. It was used to assess the participants prior knowledge. The tasks were presented starting from type 0 and difficulty 0 with alternating type and increasing difficulty to type 1 and difficulty 5. The presented tasks in the pretest were identical for every learner.

Each of the three interim tests consisted of six tasks, one of each of the levels of difficulty 1, 3, and 5 of each type. As every interim test was followed by a motivational questionnaire, and we wanted to influence self-assessment of motivation as little as possible through preceding task difficulty, we finished every interim test with a medium level of difficulty, i.e. difficulty 3. Therefore, the sequence of tasks was difficulty 5, difficulty 1, and difficulty 3 with alternating types. As in the pretest, the presented tasks in each interim test were identical for every learner. Different interim tests, however, presented different tasks.

The first learning phase started with a randomly chosen task of type 0 and difficulty 0. Using the Simple Adaptation, as explained below, this enabled the learner to quickly reach an adequate basic level of difficulty, which served as a basis for the Motivation-Based Adaptation. The following tasks were randomly chosen from the pool of learning tasks with the required type and difficulty. Every other learner was presented with a different sequence of tasks in their learning phases. The first task of a new

learning phase was adapted with respect to the last task of the last learning phase. Every task was presented no more than once, unless the pool of tasks of a specific type and level of difficulty held no more unused tasks. In this case, an already presented task was reused. Tasks from the test phases were not presented in the learning phase.

As stated above, the types of task were not adapted, but were instead alternated, to give the learner some variety. For adaptation, we concentrated on the level of difficulty. We limited the choice to three actions we considered appropriate for a learning program: increase the level of difficulty by one, decrease the level of difficulty by one, or leave the level of difficulty unchanged. We assumed that an increase or decrease of difficulty by more than one would lead to a change being too large for adaptation.

The Simple Adaptation (SA) was applied within a learning phase. It reacted to the correctness of the learner's answers and adapted the level of difficulty if the last two or more tasks were of the same level of difficulty. The adaptation was applied as follows: If the learner solved the last two tasks correctly, we assumed that they mastered the current difficulty, and the level was increased. If the learner could not solve at least one of the last two tasks correctly, we assumed that the current difficulty exceeded the learner's abilities, and the level was decreased. If the learner solved one of the tasks correctly, no change was applied to the level of difficulty. This adaptation follows the ZPD concept in an attempt to avoid the negative impact of the expertise reversal effect. This means that we tried to adapt the level of difficulty to match the user's current expertise.

The focus of this study was on Motivation-Based Adaptation (MBA). This adaptation of task difficulty was applied at the beginning of each learning phase. The objective of our work was to find plausible rules for MBA. These adaptation rules should be based on the motivation reported in the questionnaire given before the respective learning phase, and should maximise learning success.

Learning success was calculated from the interim test following a learning phase. Prior knowledge was excluded, using the interim test preceding the learning phase. More precisely, we defined learning success as the difference in ratios of correct answers in the two interim tests. Consequently, learning success could range from -1 to 1.

In our study MBA was applied randomly. That is the level of difficulty was randomly increased, decreased, or not changed according to the motivational states. This provided us with a broad sample of motivational states and adaptations from which we could extract those rules that promised the best learning success. Notice that a proposed increase (or decrease) in difficulty might not be possible because the highest (or lowest) level had already been reached. In this case, the e-learning program left the level of difficulty unchanged.

Information about a learner's performance during the program was stored in a file. This information included the learner's answers to the motivational questionnaires and their results for every task presented. The stored results of a presented task consisted of the task's type and level of difficulty, the correctness of the learner's answer and the time the learner needed to solve the task. Of the stored information, the four motivational factors and the MBA of each learning phase were independent variables.

The dependent variable was the resulting learning success, which was computed from the test phases before and after the corresponding learning phase.

Procedure and scoring

The study took place on three sequential days with two to four groups per day and one to six participants per group. Six PCs running *Windows XP* were used to run the e-learning program ELMA, which was implemented using Java, and to collect the data. Participants could work through the program in their own pace, apart from the 60-second time limit to solve a single task. The participants were also asked to fill in an additional questionnaire, which collected demographic data and comments. This data was tied to the participants' learning data but otherwise was anonymous.

When the participants started the program, they were presented first with several text pages that gave an introduction to the program and the presented tasks. They were then given the pretest and the motivational questionnaire, which measured the learner's prior knowledge and current motivation, respectively. The first learning phase started with 10 tasks which were adapted according to the SA scheme, followed by the first interim test. Two learning phases formed the key part of the program. Each consisted of the shortened motivational questionnaire to assess the learner's current motivation, the learning phase of 10 tasks, and the interim test to measure learning success. At the end of the program, the complete motivational questionnaire was presented once more to gather more information about the learner's motivational state. The results were saved, and the learners were thanked for their participation on a last text page.

Results

All but one participant stayed within two times the standard deviation of the mean value for the average time needed to solve a single task and the percentage of correctly solved tasks. One participant took a shorter average time to solve a single task while showing a lower percentage of correct answers. This correlation can be expected and may stem from low motivation. Therefore, we did not consider these values outliers and kept them in the data set.

Furthermore, no participant was able to solve more than 82% of the tasks. Similarly, no participant gave less than 31% correct answers, with the probability of solving a task correctly by randomly choosing an answer being 20%. This suggests that the overall difficulty of the tasks was adequate for the learners and floor and ceiling effects were likely to be negligible.

Learning success ranged from -0.5 (in one case) to +0.5 (in three cases), with the theoretically possible learning success ranging from -1 to 1. In other words, the learning success showed a high degree of variability for different adaptations for different learners in different situations. This suggests that finding an appropriate adaptation for a specific learner in a specific situation may indeed have a large influence on the learning success. The average learning success was 0.0045 in the first learning phase and 0.144 in the second phase. This means that on average, learners only showed a very small improvement, with roughly 0.85 more correct answers in the test after the second learning phase compared to the test before the second learning phase. In the first learning phase, average learning success was even smaller.

Table 1 shows the average value and the standard deviation for each of the motivational factors and the overall motivation, as they were reported by the participants in the complete motivational questionnaire at the beginning (Q0) and the end (Q1) of the program. Motivation for each factor could range from 0 (very low) to 6 (very high). In both questionnaires *anxiety* shows the lowest average value and interest the highest.

Table 1: Average value and standard deviation for each motivational factor in the two complete motivational questionnaires at beginning (Q0) and end (Q1) of the program

Motivational factor	Motivational questionnaire			
	Q0		Q1	
	M	*SD*	*M*	*SD*
Anxiety	2.4	1.5	2.2	1.5
Probability of success	3.5	1.3	3.5	1.3
Interest	3.6	1.2	3.6	1.2
Challenge	2.8	1.7	2.8	1.3
Overall	3.1	1.4	3.0	1.5

From the start to the end of the program, *anxiety* decreased slightly on average, while the other factors and overall motivation hardly showed any changes. Cohen's d suggests that motivation did not change considerably in the course of the study (d = 0.033 overall, d = 0.14 for *anxiety*, and d < 0.014 for the *interest*, *challenge*, and *probability of success*).

A decrease in difficulty led to the highest learning success on average, while an increase in difficulty led to the lowest learning success on average. The F-value for these adaptations is F = 0.468, and the p-value is p = 0.63. Overall, adaptation independent of motivation did not significantly influence learning success.

Table 2: An overview on the distribution of motivation and MBA in the study

Anxiety	Interest	Low challenge			Medium challenge			High challenge		
		Low success	Med. success	High success	Low success	Med. success	High success	Low success	Med. success	High success
Low	Low	1,0,0	3,0,0	2,0,0	0,0,0	2,2,1	1,1,1	0,0,0	0,1,0	1,2,0
	Med.	1,0,0	1,1,0	0,0,0	0,0,1	0,0,0	0,0,0	0,0,0	0,0,0	1,3,2
	High	0,0,0	0,0,0	1,2,1	0,0,0	0,0,0	0,2,0	0,1,0	0,0,0	0,0,0
Med.	Low	0,0,3	0,1,0	0,0,0	0,1,0	1,3,0	0,0,0	0,0,0	0,0,0	0,1,1
	Med.	0,0,0	0,0,0	0,0,0	0,1,0	1,1,0	0,0,0	0,0,0	0,0,1	0,1,0
	High	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,1,0	1,0,0
High	Low	2,0,0	0,0,0	0,0,0	0,0,1	0,1,0	0,0,0	0,0,0	0,1,0	0,0,0
	Med.	1,0,0	0,0,0	0,0,0	0,0,0	0,1,1	0,0,0	0,0,2	1,1,1	1,0,0
	High	0,0,0	1,0,0	0,0,0	0,0,0	0,1,0	0,0,0	0,0,0	1,1,2	0,1,0

Table 2 shows the distribution of motivation and adaptation in the study. Every table entry shows the amount of data collected for this motivational state, two sets per participant from two learning phases, further sorted by adaptation. Each participant may report up to two different motivational states or receive different adaptations, one in each learning phase. Similarly, two different participants may report the same motivational state and/or receive the same adaptation. Each motivational factor is divided into three states, low, medium ('med.'), and high. The four motivational states

are anxiety, challenge, interest and probability of success. For example, the entry in the top left corner of the figure shows the amount of data collected where the participant reported all of the motivational factors as low. The next entry to the right shows the amount of data collected where *anxiety*, *interest*, and *challenge* were low, and *probability of success* was medium, and so on. For example, a table entry 1,2,3 means that for this motivational state the difficulty was decreased once, not changed twice, and increased three times in the course of the study. Motivational states that are not covered by the data, i.e. were not assessed by any participant, are highlighted.

Table 2 gives information about two aspects. First, the parts of the problem space covered by the collected data are identified. Second, it shows which motivational factors typically occurred or did not occur together during the study. For example, medium or high *anxiety* with simultaneously low *challenge* and medium or high *probability of success* was only reported once. Similarly, high *probability of success* only occurred together with medium or high *anxiety* if *challenge* was high. The most frequent motivational state was low *anxiety* and medium *interest* with high *challenge* and high *probability of success*. This motivational state was measured six times in the course of the study.

The learning classifier system XCS

The study generated motivation-dependent learning success data for a widely distributed range of motivational states. To investigate if the adaptations of the level of difficulty after the motivational questionnaires influenced learning success, and if suitable adaptation strategies for e-learning environments can be derived, the data collected in the study had to be analysed thoroughly. The learning classifier system XCS was used to interpret the data and to extract rules for motivation-based adaptation. XCS was introduced by Wilson (1995) and has been applied successfully to various classification, data mining, and reinforcement learning problems. The system optimises a population of classifiers, or rules, by a reinforcement learning (RL) -based rule evaluation mechanism, and an evolutionary-based rule structure optimization mechanism. The rule evaluation mechanism updates the reward prediction, the average error and the experience of a rule, based on the rule's performance with respect to the sampled input data. The rule evolution system produces new rules based on an accuracy-based fitness signal, which is derived from the RL mechanism, to search the problem space for an accurate and general representation of the problem underlying the structure.

The data from which XCS learned consisted of the four motivational factors reported by the participant at the beginning of each learning phase, the applied adaptation of the level of difficulty, and the participant's learning success for the respective phase. The four motivational factors could each take one of three possible values. These values represented low, medium or high motivational states. The level of difficulty was decreased, maintained or increased by one level. The learning success was determined by the ratio of correct answers in the interim test following the current learning phase minus the ratio of correct answers in the interim test preceding the current learning phase. XCS learns iteratively, sampling the 74 available data points randomly, with replacement in our case. Each rule predicts the expected learning success for a certain change in level of difficulty – or a certain range of changes – given the reported motivation of the student.

To ensure the suitability of XCS for extracting rules from this very sparse and presumably highly noisy data set, the system was thoroughly tested on artificial data with similar sparsity and expected noise levels (Endler, Butz & Rey, 2011) The results verified the suitability of XCS for structure extraction from such data sets.

Extracted rules

Rules were extracted from ten runs of XCS with the available data, each of which achieved an acceptable average error and a reasonable number of rules. From these ten batches of rules, only those rules with high fitness and experience were considered further. Of these remaining rules, those that appeared in at least two different runs were extracted. The resulting final batch of twelve rules is shown in Table 3.

Table 3: Rules derived from the study's data using the learning classifier system XCS

No.	Anxiety	Success	Interest	Challenge	Adaptation	Learning success
1	0-1	0-1	0-1	1-2	0	0.227, 0.333
2	0-2	2	0-1	0-1	0-1	-0.121, -0,104, -0.077, 0.097
3	0-2	1-2	1	0-2	0-1	0.096, 0.133
4	0-2	0-2	1	0-2	0-1	0.120, 0.333
5	1-2	0-2	0-1	0	0-2	-0.021, 0.038
6	1-2	1-2	0-1	1-2	0-2	0.170, 0.125, 0.134
7	1-2	0-2	0-2	0	0-2	-0.046, -0.027, 0.035
8	2	0-2	0-2	0-2	1	0.204, 0.209, 0.241, 0.249, 0.263
9	0	0-1	0-2	0-2	1-2	-0.038, -0.027, -0.003, 0.002
10	0-2	2	1-2	0-2	1-2	-0.046, 0.017, 0.094
11	0-2	2	1-2	1-2	1-2	-0.034, 0.013
12	0-2	2	2	0-2	1-2	-0.063, -0.019

The first column gives a number to every rule. The next five columns show the rule, starting with the four motivational factors *anxiety*, *probability of success* (success), *interest* and *challenge*, where 0 means low, 1 means medium, and 2 means high. This is followed by the proposed adaptation, either decrease by one (0), no change of difficulty (1), or increase by one (2). Motivational factors and adaptation can also take ranges of their domain. The last column states the average expected learning success per run for the proposed adaptation, or any one of the proposed adaptations, given a motivational state covered by the range of states as specified by the rule. As the same rule may not appear in each run, the number of learning success predictions varies from two to five. Learning success predictions for the same conditions always show a similar value for each run with the highest deviation being 0.218 for Rule 2. Furthermore, Rule 4 subsumes Rule 3, Rule 7 subsumes Rule 5, Rule 10 subsumes Rule 11 and Rule 12. Rules 5, 6 and 7 give no indication for the best adaptation, but do rate a certain combination of motivational factors and their influence on learning success.

Discussion

The objective of this work was to find appropriate adaptation strategies for MBA in a simple e-learning environment. To achieve this, we used an LCS to extract general

rules from the results of our study. In the following, we discuss if these rules are plausible in the light of available theories on motivation and learning.

ZPD and flow

We interpret the *challenge* variable as an indication for the interaction between task difficulty and learner's expertise, as defined by the ZPD. This interpretation assumes that high *challenge* indicates that difficulty exceeds the learner's expertise while low *challenge* indicates that the learner's expertise exceeds difficulty. In this scenario, medium *challenge* should imply a balance of difficulty and expertise.

High *challenge* should then suggest that difficulty is too high, and the learner will leave the ZPD. In this case, reducing the level of difficulty will return the learner into the ZPD and therefore, usually increase learning success. This is confirmed by Rule 1. If, on the other hand, difficulty is decreased further when a learner reports low to medium *challenge* they leave the ZPD because difficulty is too low for their expertise, resulting in low learning success, as seen in Rule 2. The same holds true if difficulty is not decreased for high *challenge,* as seen in Rule 11. All other rules give no indication for either *challenge* or adaptation and are therefore not applicable to this theory.

The theory on flow suggests that high difficulty for low expertise will lead to anxiety, while low difficulty for high expertise will lead to boredom. Our rules cannot support that too high a difficulty will lead to anxiety, but it seems plausible that high *probability of success* combined with low *interest* might indicate boredom. We can observe this combination of motivational factors in Rule 2. Boredom may indicate that difficulty is too low, i.e. the learner has left the ZPD. This is also indicated by low *challenge* values. Consequently, further decrease or no change of difficulty should result in a low learning success, as predicted by Rule 2.

Yerkes-Dodson Law

Anxiety is a candidate to be an activating motivational factor, as defined by the Yerkes-Dodson Law. If learners show no anxiety, they may not see the need to concentrate on the task. If, on the other hand, anxiety is too high, they may not be able to concentrate on the task. In the study, self-assessed *anxiety* must be seen in the context of performance as having no consequences for the participants, as they are not graded or otherwise judged for their results. Furthermore, the simple adaptation within a learning phase should keep the difficulty of the tasks at a medium level, as adjusted to a specific learner's abilities. Therefore, we assume that even high reported *anxiety* does not leave the range where it is activating, rather than blocking, learning. Consequently, high (or low) *anxiety* should lead to a high (or low) learning success. This is confirmed by rules 8 and 9, respectively. Medium to high *anxiety*, which can be found in rules 5, 6, and 7, shows a wider range in learning success but mainly within the boundaries of the learning success of rules 8 and 9.

Interest can be analysed using the same law. *Interest*, however, may take the full range in this study so that we expect medium *interest* to result in high learning success, and low or high *interest* in a lower learning success. Rules 3 and 4 show an acceptable learning success for medium *interest*. Rules 2 and 5 indicate a low learning success for low to medium *interest*, and rules 10, 11, and 12 show a low learning success for medium to high *interest*. Rules 1 and 6, however, contradict this assumption. Rule 1

shows a very high learning success for low to medium *interest*. This might be due to other factors such as the ZPD explained above. Rule 6 also shows a rather high learning success for low to medium *interest*.

Rule 6 differs from Rule 5 mainly in *challenge*, suggesting that high *challenge* may compensate for the lack of interest. Looking at the other rules, low to medium *interest* together with high *challenge* always promises a better learning success than low to medium *interest* with low *challenge* (see rules 1, 6 and 2, 5, respectively).

Considered by itself, *challenge* shows a similar behaviour as *anxiety,* if we assume that challenge reported in the context of the study will not exceed the activating range, because all tasks are still very simple in the sense that there are no complex contexts for the learner to take into account. Rules 2, 5 and 7 predict a low learning success for low or low to medium *challenge* while rules 1 and 6 predict a high learning success for medium to high *challenge*. Rule 11, however, does not support this theory, predicting a rather low learning success for medium to high *challenge*.

Overall, the extracted rules can largely be embedded into existing theories on motivation and learning. Particularly, the observed influence of *challenge*, interpreted as an indication for the interaction of task difficulty and expertise, and of adaptation on learning strongly supports the concept of the ZPD. Furthermore, *anxiety* and *interest* seem to show an activating effect as defined by the Yerkes-Dodson Law.

Limitations

Several limits of the study, due mainly to restrictions of time and resources, should be taken into account when appraising the results of this work. With 37 participants, the amount of collected data was rather low, meaning that the entire problem space could not be covered. Similarly, only a small number of values could be obtained for a single motivational state. However, we believe that the available data was sufficient to detect some general rules, which indicate the influence of motivation on learning success and provide a solid basis for motivation-based adaptation strategies. However, further evaluation using a larger amount of data will be necessary to empirically investigate the quality and performance of the extracted rules.

Furthermore, more than 50% of the participants were students. Thus, the group of participants did not form a representative cross-section of the population. Consequently, the results do not necessarily apply in the same way for the general population. They can also not be generalised for different types of tasks since the tasks used in the study asked for mathematical abilities. Whether the motivation-based rules derived here are applicable to other types of learning tasks and other population groups needs to be assessed in future research.

Another critical point is the self-assessment of motivation. While self-assessment is easy to measure and has been used in a number of studies (e.g. Astleitner & Koller, 2006; Martens, Gulikers & Bastiaens, 2004; Song & Keller, 2001), there is no guarantee that a learner's reported motivational state does, in fact, reflect their true motivational state. They may not be able to evaluate their own motivation reliably, or they may report the motivation they think they are expected to report. However, as the validity of the questionnaire used in this work is supported by several studies (see Rheinberg

et al. 2001), and there is to date no provably more accurate or natural way to measure motivation, we believe that we chose the most appropriate method available.

Future directions

Due to the small amount of data available, the utility of the extracted rules can only be proven preliminarily. Nonetheless, we were able to develop a well-grounded hypothesis on adequate motivational adaptation and its influence on learning success. Future work will have to evaluate the quality of the extracted rules and motivation-based adaptation strategies derived from them. Another study, using ELMA with these strategies, will be necessary to achieve this. This study will adapt difficulty based on motivation, following the extracted rules or a subset of these rules. Comparison with a Yoked Design control group should tell us if the proposed adaptation rules promise an increased learning success, and if our approach for motivation-based adaptation can help students learn more efficiently.

With further research in this field and better, more exact ways to measure motivation, we believe that this work can help lay the groundwork for motivation-based adaptive E-learning environments. Eventually, e-learning may be adapted to the individual student needs and state of minds, and thus become more effective and motivating in all e-learning areas, including simple vocabulary trainers and even entire university programs.

References

Astleitner, H. & Koller, M. (2006). An aptitude-treatment-interaction-approach on motivation and student's self-regulated multimedia-based learning. *Interactive Educational Multimedia,* 13, 11-23. http://www.raco.cat/index.php/iem/article/viewFile/205273/273808

Broadhurst, P. L. (1957). The interaction of task difficulty and emotion: The Yerkes-Dodson Law revived. *Acta Psychologica,* 16, 321-338. http://dx.doi.org/10.1016/0001-6918(59)90105-2

Butz, M. V. (2006). *Rule-based evolutionary online learning systems.* Heidelberg: Springer.

Chen, C.-M., Lee, H.-M. & Chen, Y.-H. (2005). Personalized e-learning system using item response theory. *Computers & Education*, 44(3), 237-255. http://dx.doi.org/10.1016/j.compedu.2004.01.006

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

Endler, A., Butz, M. V. & Rey, G. D. (2011). Extracting adaptation strategies for e-learning programs with XCS. *In GECCO '11: Proceedings of the 13th annual conference companion on Genetic and Evolutionary Computation*, pp.743-746. New York: ACM. http://dx.doi.org/10.1145/2001858.2002081

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning.* Reading, MA: Addison-Wesley.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509-539. http://dx.doi.org/10.1007/s10648-007-9054-3

Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23-31. http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1141&context=edupapers

Keeley, J., Zayac, R. & Correia, C. (2008). Curvilinear relationships between statistics anxiety and performance among undergraduate students: Evidence for optimal anxiety. *Statistics Education Research Journal, 7*(1), 4-15. http://www.stat.auckland.ac.nz/~iase/serj/SERJ7(1)_Keeley.pdf

Keller, J. M (2006). *What is motivational design?* Florida: Florida State University. http://www.arcsmodel.com/pdf/Motivational%20Design%20Rev%20060620.pdf

Kim, K.-J. & Frick, T. W. (2011). Changes in student motivation during online learning. *Journal of Educational Computing Research*, 44(1), 1-23. http://dx.doi.org/10.2190/EC.44.1.a

Heller, K. A. & Perleth, C. (2000). *KFT 4-12+R Kognitiver Fähigkeitstest für 4. bis 12 Klassen, Revision.* Göttingen: Hogrefe.

Holland, J. H. (1975). *Adaptation in natural and artificial systems.* Ann Arbor, MI: University of Michigan Press.

Martens, R. L., Gulikers, J. & Bastiaens, T. (2004). The impact of intrinsic motivation on e-learning in authentic computer tasks. *Journal of Computer Assisted Learning*, 20, 368-376. http://dx.doi.org/10.1111/j.1365-2729.2004.00096.x

Mödritscher, F., Garcia-Barrios, V. M. & Gütl, C. (2004). The past, the present and the future of adaptive e-Learning. *Proceedings of the International Conference on Interactive Computer Aided Learning (ICL2004).* http://www.iicm.tugraz.at/www/Research/Publications/2004/_id188a8_/adaptive_e-learning/adaptiv_e-learning.pdf

Rey, G. D. (2009). *E-Learning: Theorien, Gestaltungsempfehlungen und Forschung*. Bern: Huber. http://www.elearning-psychologie.de/

Rey, G. D. & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied,* 17, 33-48. http://psycnet.apa.org/journals/xap/17/1/33/

Rheinberg, F., Vollmeyer, R. & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47, 57-66. http://www.psych.uni-potsdam.de/people/rheinberg/messverfahren/FAMLangfassung.pdf

Rudolph, U. (2009). *Motivationspsychologie kompakt.* Weinheim: Beltz.

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C. & Bigdeli, A. (2008). "How do you know that I don't understand?" A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24, 1342-1363. http://dx.doi.org/10.1016/j.chb.2007.07.008

Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development.* (pp. 183-212) Hillsdale, NJ: Erlbaum.

Shute, V. & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist*, 38(2), 105-114. http://dx.doi.org/10.1207/S15326985EP3802_5

del Soldato, T. & du Boulay, B. (1995). Implementation of motivational tactics in tutoring systems. *Journal of Artificial Intelligence in Education*, 6(4), 337-378. http://dl.acm.org/citation.cfm?id=225657&CFID=130865741&CFTOKEN=56626276

Song, S. H. & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. Educational Technology Research and Development, 49(2), 5-22. http://dx.doi.org/10.1007/BF02504925

Sutton, R. S. & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

De Vicente, A. (2003). *Towards tutoring systems that detect students' motivation: An investigation.* PhD thesis, School of Informatics, University of Edinburgh, UK. http://www.iac.es/galeria/angelv/papers/thesis.pdf.gz

Vygotski, L. S. (1963). Learning and mental development at school age. In B. Simon & J. Simon (Eds.), *Educational psychology in the USSR.* (pp. 21-34). London: Routledge & Kegan Paul.

Wilson, S. W. (1995). Classifier fitness based in accuracy. *Evolutionary Computation*, 3(2), 149-175. http://dx.doi.org/10.1162/evco.1995.3.2.149, also at http://www.eskimo.com/~wilson/ps/xcs.pdf

Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology,* 18, 459-482. http://psychclassics.yorku.ca/Yerkes/Law/

Authors: Anke Endler, Department Psychology III, Institute for Psychology
Julius-Maximilians-Universität Würzburg, Germany.
Email: anke.endler@uni-wuerzburg.de
Web: http://www.coboslab.psychologie.uni-wuerzburg.de/people/anke_endler/

Dr Günter Daniel Rey *PhD* (corresponding author)
Department Educational Psychology, Institute for Psychology
FernUniversität in Hagen, Germany. Email: guenter-daniel.rey@fernuni-hagen.de
Web: http://www.fernuni-hagen.de/universitaet/stimmen/rey.shtml

Prof Dr Martin V. Butz, Department of Computer Science, Faculty of Science,
University of Tübingen, Germany. Email: butz@informatik.uni-tuebingen.de
Web: http://www.wsi.uni-tuebingen.de/lehrstuehle/cognitive-modeling/staff/staff/martin-v-butz.html