Renkl, A., Skuballa, I. T., Schwonke, R., Harr, N., & Leber, J. (2015). The Effects of Rapid Assessments and Adaptive Restudy Prompts in Multimedia Learning. *Educational Technology & Society, 18* (4), 185–198.

The Effects of Rapid Assessments and Adaptive Restudy Prompts in Multimedia Learning

Alexander Renkl*, Irene T. Skuballa, Rolf Schwonke, Nora Harr and Jasmin Leber
Department of Psychology, University of Freiburg (Germany), Freiburg, Germany // renkl@psychologie.uni-freiburg.de // irene.skuballa@uni-tuebingen.de // schwonke@psychologie.uni-freiburg.de // harr@psychologie.uni-freiburg.de // jasmin.leber@psychologie.uni-freiburg.de
*Corresponding author

ABSTRACT

We investigated the effects of rapid assessment tasks and different adaptive restudy prompts in multimedia learning. The adaptivity was based on rapid assessment tasks that were interspersed throughout a multimedia learning environment. In Experiment 1 (N = 52 university students), we analyzed to which extent rapid assessment tasks were reactive (i.e., whether these tasks change what should be measured) and, thus, had per se positive effects on learning. In Experiment 2 (N = 41 university students), we analyzed the advantages and disadvantages of specific and unspecific restudy prompts (i.e., focus on a very specific piece of knowledge or on the corresponding knowledge sub-area). We found no reactivity associated with the assessment tasks. Most specific knowledge gaps could be closed by either type of prompt. However, unspecific prompts fostered the overall learning outcomes more than specific prompts. The present adaptation procedure is a good starting point for developing powerful adaptation mechanisms.

Keywords

Rapid assessment, Adaptive prompts, Multimedia learning, Specificity of prompts

Introduction

Educational technology provides us with learning environments which adapt learning paths to the individual's needs. Students learn by problem solving via a number of well-established systems. Their corresponding attempts can be used to adapt hints and learning tasks to their individual needs (e.g., Cognitive Tutor; http://www.carnegielearning.com; Koedinger & Corbett, 2006). Other environments provide expository instruction and present multimedia contents to learners. In these cases, explicit probes or diagnostic tasks must be interspersed for adaptation purposes (e.g., SmartBook; http://www.engadget.com/2013/01/08/mcgraw-hill-smartbook). An efficient means of such diagnosis is the use of rapid assessment procedures as developed by Kalyuga and colleagues within the framework of cognitive load theory (e.g., Kalyuga, 2008; Kalyuga & Sweller, 2005). Rapid assessment tasks can be used to detect learners' knowledge gaps and to adapt further learning paths to these deficits.

We conducted two experiments on the effects of rapid assessment tasks and two types of adaptive restudy prompts in multimedia learning. Rapid assessments are tasks that should be fulfilled quickly and that are interspersed throughout a learning environment. The first experiment addressed the extent to which rapid assessment tasks can be regarded as a non-reactive diagnostic method, meaning that they do not change or influence what they measure (i.e., knowledge states). Note that in the first experiment, rapid assessment was not used to adapt instruction; we tested just its potential reactivity. In the second experiment, rapid assessment was used to diagnose knowledge gaps and adapt instruction. Specifically, we tested the effects of two types of restudy prompts triggered by wrong responses to rapid assessment tasks.

Adaptive learning systems

If you confront different learners with a learning environment, they are sure to differ in their learning outcomes (e.g., Ackerman & Lohman, 2006). These differences can be attributed to the learners' varying prerequisites and the fact that a one-size-fits-all environment cannot be optimal for different learning prerequisites. One remedy is to use adaptive learning environments (e.g., Shute & Zapato-Rivera, 2008; Vandewaetere, Desmet, & Clarebout, 2011). Adaptation can refer to different sizes of grain in this context. At its coarsest, macro-adaptation refers to assigning different learning environments to different learners (Park & Lee, 2003). When the grain is fine, micro-adaptation refers to adapting instructional events during learning to a learner's cognitive or affective states. We focused on micro-adaptation in our experiments.

Adaptive systems can react to different learner characteristics such as (prior) knowledge states, working memory capacity, cognitive styles, motivation, or emotional states (see, e.g., Vandewaetere et al., 2011). Knowledge-related variables have most frequently been used to adapt instruction (e.g., prior knowledge, knowledge states, or identified knowledge gaps). This emphasis on knowledge-related factors is not surprising, given their conceptual affinity with the knowledge-related learning goals of most adaptive systems. Furthermore, knowledge prerequisites are the most important factor for further learning both positively and negatively. Correct prior knowledge is usually the most important factor facilitating further learning (e.g., Dochy, de Rijdt, & Dyck, 2002; Kalyuga, 2012). Incorrect knowledge (e.g., misconceptions or misunderstanding) is usually the most substantial barrier for further learning (Ambrose & Lovett, 2014). In addition, research on aptitude-treatment interactions and on the expertise-reversal effect (i.e., sensible instructional features for novices lose their effectiveness with more knowledgeable learners) has clearly shown that different instructional procedures are best used for learners with different knowledge states (for an overview see Lee & Kalyuga, 2014). It goes without saying that knowledge-related indicators have often been used in adaptive systems.

Rapid assessment

Within a cognitive load framework, Kalyuga and colleagues (e.g., Kalyuga & Renkl, 2010) have produced many findings on the expertise-reversal effect. This effect means that learners with different (prior) knowledge levels profit from different instructional features (see Lee & Kalyuga, 2014). In this context, Kalyuga and colleagues (e.g., Kalyuga, 2006, 2008; Kalyuga & Sweller, 2005) developed a rapid assessment method that enables the online-diagnosis of knowledge states and deficits. For this purpose, (small) tasks, which should be answered rapidly, are interspersed throughout a learning environment. Some formats of rapid assessment tasks can be sensibly applied just in procedural domains (e.g., the first-step method that requires learners to provide a first solution step to a problem). A format applicable in most domains is rapid verification. Rapid verification tasks present solution steps or statements, and the learners must quickly determine whether the presented information is right or wrong (see also Roelle, Berthold, & Renkl, 2014).

Kalyuga and Sweller (2004) investigated the effects of adapting instruction on the basis of rapid assessment. They exploited the finding that with a learner's growing knowledge level, subsequent learning tasks are best for acquiring cognitive skills: first fully-worked examples, then faded-worked examples (i.e., worked examples containing some gaps to be filled in), and finally problem solving tasks (e.g., Renkl, 2014; Renkl & Atkinson, 2003). In the adaptation condition, learners studied fully-worked examples, faded-worked examples, or problem solving exercises depending on their prior knowledge level and their progress during the learning phase, as assessed by rapid assessment tasks. In the control group, each learner followed the learning path of a yoked learner from the adaptation group. There were no rapid assessment tasks in the control group to monitor learners' progress during the learning phase. Kalyuga and Sweller (2004) observed superior learning outcomes in the adaptation group. However, it is not totally clear to what extent this advantage was due only to the adaptive presentation of learning tasks or to working on the additional test tasks as well. The latter activity can enhance learning outcomes, as research on the testing effect has revealed (Rawson & Dunlosky, 2012; Roediger, Putnam, & Smith, 2011). Against this background, it would be interesting to test whether just working on rapid assessment tasks already exerts positive effects on learning outcomes.

Study prompts

Prompts are questions or hints provided to learners. They are designed to induce productive processing of learning materials (e.g., Devolder, van Braak, & Tondeur, 2012; Pressley et al., 1992). The type of processing that is induced depends on its specific purpose. For example, prompts were successfully employed to foster self-explanations (Chi, de Leeuw, Chiu, & LaVancher, 1994), comparison of cases (Gentner, Loewenstein, & Thompson, 2003), or self-regulation activities (Bannert & Reimann, 2012). In some learning environments, the prompts are adaptive in the sense that they depend on the learners' prior behavior (e.g., Graesser, Jeon, & Dufty, 2008; Nückles, Hübner, Dümer, & Renkl, 2010).

Although there are several studies comparing different types of prompts (e.g., Ifenthaler, 2012; Nückles, Hübner, & Renkl, 2009), there is hardly any evidence as to which type of prompts are best used for closing knowledge gaps (see also Devolder et al., 2012). In some cases, learners' incorrect answers might indicate that a very specific piece of knowledge is missing (e.g., the fact that it is the nucleus where DNA doubles during mitosis). In other cases, the corresponding sub-area of the learning contents might be missing as well (e.g., what happens in general in the nucleus during mitosis). If a prompt just encourages learners to close the very specific knowledge gap, the potentially underlying problem (that the entire sub-area is unknown) goes unaddressed. In contrast, if a prompt encourages not just looking up the specific missing piece of knowledge, but considering the knowledge sub-area as well, broader effects on learning could be expected. On the other hand, unspecific prompts might be less efficient than specific prompts when just such a specific knowledge gap should be closed. In addition, unspecific prompts might induce unnecessary processing of already understood materials (see the *redundancy effect* in cognitive load theory; Sweller, Ayres, & Kalyuga, 2011).

Overview of the present experiments

We conducted two experiments. Each investigated one main element of our adaptation procedure (i.e., rapid assessment-based provision of restudy prompts). In Experiment 1, we analyzed to which extent rapid assessment tasks have per se positive effects on learning outcomes. In addition, we considered factors that might contribute to such an effect (e.g., rapid assessment tasks might motivate deeper processing on the following materials). In Experiment 2, we analyzed the advantages and disadvantages of different types of restudy prompts that varied in their focus (i.e., on a very specific piece of knowledge or on the corresponding knowledge sub-area). The specific hypotheses tested in these two experiments are presented right before the corresponding method sections.

Experiment 1

We investigated whether adding diagnostic tasks to a learning environment is reactive in the sense that these tasks “alone” already foster knowledge acquisition. More specifically, we tested the potential effect of rapid assessment tasks on learning outcomes (*Learning-Outcomes Hypothesis*). Findings on the testing effect suggest that there might be such an effect (e.g., Rawson & Dunlosky, 2012; Roediger et al., 2011). As the assessment tasks also familiarize students with test tasks from the learning domain, we expected that the learners with rapid assessment tasks would perceive the problems in the test on learning outcomes (posttest) as easier and requiring less mental effort (*Subjective-Load Hypothesis*).

In addition, we explored a number of potential mediational mechanisms that may account for a learning effect: The rapid assessment tasks may orient the learners about the learning goals (see Roediger et al., 2011) (*Orientation Hypothesis*). The expectation of up-coming test tasks may motivate longer (*Learning-Time Hypothesis*) and deeper processing of the materials (Roediger et al., 2011) (*Reflection Hypothesis*). Providing learners with the opportunity to check what they have learned may make them perceive the learning environment as more interesting and useful. In addition, if rapid assessment tasks help learners comprehend the learning contents, this factor would also contribute to situational interest (Schraw, Flowerday, & Lehmann, 2001) (*Interest Hypothesis*). We planned to test mediation effects in the case of significant effects on the variables that might explain a possible effect of rapid assessment on learning outcomes (e.g., orientation about learning goals or learning time).

Method

Sample and design

Fifty-two university students from a psychology program took part in this study (age: M = 24.63, SD = 5.65). They received study credits for participation. The students were randomly assigned to two conditions (n = 26 in each condition): computer-based learning environment about mitosis with or without rapid assessment tasks (in form of rapid verification). No feedback was provided for the assessment tasks as we wanted to test the reactivity of the tasks themselves. Note also that the assessment tasks were not used for adaptation. The central dependent measure referred to the learning outcomes and the subjective load while working on the posttest. In addition, potential mediators were assessed: orientation about the learning goals, learning time, reflection, and situational interest.

Learning environment

The learning environment topic was the process of mitosis. The contents were provided by text and pictures (Figure 1). All participants were instructed to study the learning program at their own pace. They should prepare for a final test (posttest) that was “epitomized” by two exemplary items.

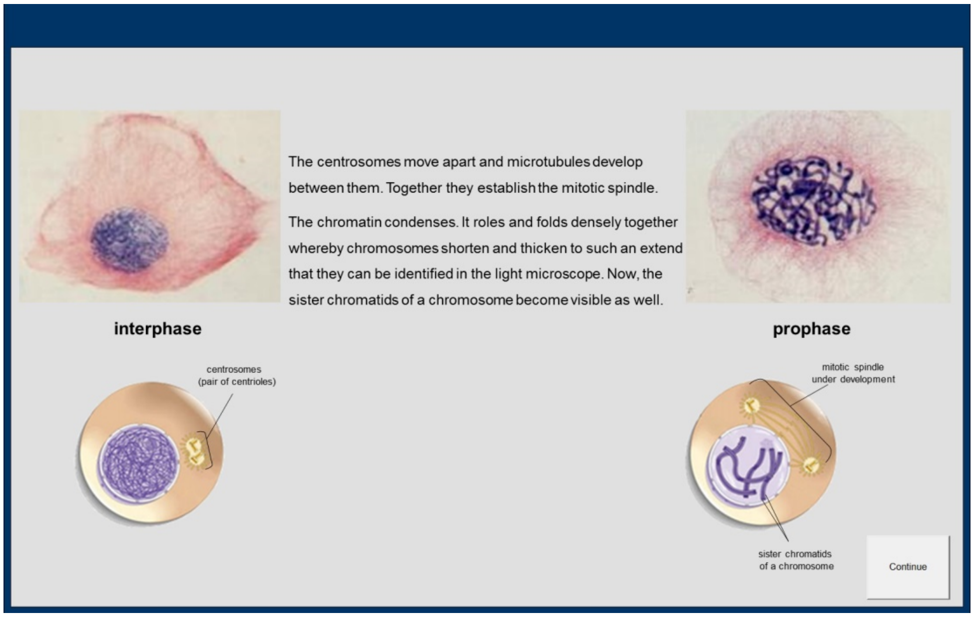

Figure 1. Screenshot from the multimedia learning environment on mitosis (Experiment 1)

The learning environment consisted of (a) an introduction including a general overview of the cell structure and of the different mitotic phases and (b) three blocks with two or three mitotic phases. The phases were explained and illustrated by schematic and realistic pictures. In order to emphasize the changes over the different stages of mitosis, the single pages simultaneously provided information about subsequent phases (Figure 1). After each block, the learners rated how well informed they felt about the learning goals, how interested they were, how much they reflected about the interrelations between the presented information, and how much they tried to anticipate what comes next (scale 1 to 7; 1: not at all).

In the rapid-assessment group, we presented 12 rapid verification tasks after each of the three blocks (4 tasks on cell structures, 4 tasks on processes, 4 tasks on functions; Kalyuga, Renkl, & Paas, 2010). Thus, there were 36 assessment tasks all in all. An exemplary verification task contained the following statement: “The microtubules form the mitotic spindle.” The learners had to click the “right” or “wrong” button. Half of the 36 statements were wrong. After 15 seconds the task disappeared. After 4 rapid assessment tasks, there was a question on how difficult it was to answer these tasks. However, we did not analyze data that were only available in the condition with rapid assessment further.

Learners from the rapid-assessment group worked on the diagnostic tasks after each block. Learners from the group without rapid assessment did not work on any tasks referring to the presented contents; there was no substitute for the rapid assessment tasks in this group.

Instruments

A *pretest* on prior knowledge about mitosis should primarily check the comparability of the experimental groups. It consisted of nine items (e.g., “Please write down five parts of the human cell”). For these open items, we provided 1 point for each aspect that was included in expert answers. Two raters coded the answers from 11 participants (i.e., about 20%). These raters reached high agreement, indicated by an ICC of .933; disagreements were resolved by discussion. We obtained an estimate of .77 (Cronbach’s alpha).

We assessed reflection with two items (i.e., how much they reflected about the interrelations within the presented information and how hard they tried to anticipate what comes next) that were present three times (i.e., after each rapid assessment block). As both sub-scales were highly correlated ($r = .68$, $p < .001$) we aggregated them to an overall reflection score (Cronbach's alpha: .90). Situational interest was assessed with one item ("This learning block on mitosis was interesting") that was presented three times (Cronbach's alpha: .92). Finally, one item assessed three times the extent to which the learners felt oriented about the learning goals ("I have an idea about what I should learn"; Cronbach's alpha: .88).

The *posttest* consisted of 22 items which were primarily verbal (e.g., "Please describe what happens during cytokinesis") or primarily pictorial (e.g., question about a schematic picture: "What is wrong in this schematic picture?"). Most of these items (i.e., 16) required open answers. We provided 1 point for each aspect that was included in expert answers. Two raters coded the answers from 11 participants (i.e., about 20%). These raters reached high agreement, indicated by an ICC of .969; disagreements were resolved by discussion. We obtained an estimate of internal consistency of .83 (Cronbach's alpha) for the overall posttest score (including 22 items). After each posttest item, we asked the learners to indicate on a 10-point rating scale how difficult the problem was and how much mental effort they had invested. We aggregated these ratings over the different posttest items and obtained a reliability estimate of .93 (Cronbach's alpha) for subjective difficulty and of .96 (Cronbach's alpha) for mental effort. Subjective difficulty and mental effort are both often used as measures of cognitive load (see, e.g., van Gog & Paas, 2008). However, as both measures correlated just moderately ($r = .36$, $p = .009$) we treated them separately in the following analyses (see also van Gog & Paas, 2008).

Procedure

The experimental sessions took place in small groups containing 5 to 12 participants. The students were welcomed in a computer room and instructed to press the "Start" button on their screens. The entire session was implemented on the computer. First, the students filled in a questionnaire on demographic data and worked on the pretest. Then, they received a short overview of the learning contents and two exemplary test questions that they should be able to answer after studying the learning environment. In the following, the procedure differed in certain respects between conditions. The students without rapid assessment proceeded directly to the learning environment. The students with rapid assessment received a short practice phase showing how to work with rapid assessment tasks. This phase was designed to prepare them for reacting "rapidly" even to the first assessment tasks. Afterwards, the students in the rapid-assessment condition worked on the same learning environment as the group without rapid assessment, except that there were three interspersed blocks with rapid assessment tasks. Note, however, that the learning phase was divided in three blocks for all learners because after each block we assessed interest, orientation about learning goals, and reflection. At the end of the session, all participants worked on the posttest. Finally, they received their credits for participation.

Results

For all analyses, we used an alpha level of .05. We report *d* as effect-size measure that was interpreted as follows: $d = .20$ as small effect, $d = .50$ as medium effect, and $d > .80$ as large effect (Cohen, 1988). Table 1 shows the means and standards deviations of the central variables in both conditions. Due to some missing data the sample size for different analyses varied between 50 and 52 participants.

The learners in the two conditions did not differ significantly in their prior knowledge, $t(50) = 0.54$, $p = .589$, $d = 0.15$. As the possible maximum pretest score was 38, the overall mean of 8.10 (about 23% correct) indicated a low level of prior knowledge. There were no significant differences between groups with respect to age, semesters at university, gender, German as mother tongue, or prior biology courses (all *p*s > .20). Hence, the learners in both conditions had comparable learning prerequisites.

Overall, there were 36 rapid assessment tasks. On average 24.69 ($SD = 4.19$) tasks were correctly answered. Hence, they were far from trivial but also easy enough to be answered correctly in two-thirds of the cases.

Table 1. Means (Standard deviations in brackets) of important variables for both conditions (Experiment 1)

	No rapid assessment (*n* = 26)	Rapid assessment (*n* = 26)
Prior knowledge (no. correct)	8.46 (4.19)	9.15 (4.95)
Learning outcomes (no. correct)	37.78 (10.86)	34.36 (12.05)
Mental effort (posttest scale from 1 to 7)	5.82 (1.63)	4.91 (1.80)
Difficulty (posttest; 1 to 7)	6.15 (1.43)	5.74 (1.99)
Learning time (min.)	11.78 (3.46)	11.21 (3.81)
Orientation about learning goals (1 to 7)	5.46 (1.17)	5.04 (1.30)
Reflection (1 to 7)	4.29 (1.24)	3.92 (1.35)
Situational interest (1 to 7)	4.52 (1.52)	4.17 (1.66)

We detected no significant group differences with respect to learning outcomes, $t(48) = -1.54$, $p = .297$, $d = -.30$ (Table 1). For exploratory reasons, we also checked whether the learning effect was moderated by prior knowledge (ATI effect). However, the respective interaction term did not reach the level of statistical significance, $F(1,46) = 1.09$, $p = .302$. Based on these findings, we rejected the *Learning-Outcomes Hypothesis*.

We found no significant group differences with respect to the perceived difficulty of the posttest, $t(49) = -0.85$, $p = .401$, $d = -.24$, or to mental effort during the posttest, $t(50) = -1.92$, $p = .061$, $d = -.53$. The *Subjective-Load Hypothesis* was not confirmed. Both groups spent on average a bit more than 11 minutes on the learning contents (without the time on assessment tasks); there was no significant group difference, $t(50) = -0.57$, $p = .574$, $d = -.16$. Hence, we rejected the *Learning-Time Hypothesis*. Note, however, that the rapid-assessment group spent an additional 3.12 minutes ($SD = 0.93$) on the rapid assessment tasks.

There were no significant group differences with respect to the learners' perceived orientation about learning goals, $t(50) = -1.23$, $p = .223$, $d = -.35$, to reflections on the learning contents, $t(50) = -1.01$, $p = .315$, $d = -.29$, or to situational interest, $t(50) = -0.81$, $p = .421$, $d = -.22$. Hence, we rejected the *Orientation Hypothesis*, the *Reflection Hypothesis*, and the *Interest Hypothesis*.

One research question referred to potential mediation effects. Rapid assessment might have fostered learning outcomes via more reflection, time on the learning contents, situational interest, or better orientation about learning objectives. However, as we identified no significant group differences concerning the latter variables, they cannot be regarded as mediators. Nevertheless, the relations between the potential mediators and learning outcomes should be reported: $r = .39$, $p = .006$, for learning time; $r = .36$, $p = .011$, for orientation about learning goals; $r = .39$, $p = .005$, for reflection; $r = .46$, $p = .001$, for situational interest. These significant correlations can be tentatively interpreted as evidence that the assessed variables were relevant for learning in the present context.

Discussion

The hypotheses on potential effects of rapid assessment were not confirmed. This is good news in the present case, as we wanted to test whether rapid assessment could be regarded as a non-reactive – or just minimally reactive – diagnostic procedure. Against the background of our results, rapid assessment does not seem to be reactive.

A possible objection against interpreting the present findings as indicating the absence of reactivity might be that the reactivity effects might have just failed to reach the level of statistical significance (e.g., because of a lack of statistical power). Note, however, that there were not only "missing" significant effects, but also descriptive mean differences that in most cases were in favor of the group without rapid assessment (Table 1). Only the perceived difficulty and mental effort with regard to the posttest were descriptively lower in the rapid-assessment condition. Hence, there was hardly any indication for reactivity. Another alternative explanation for the missing effects could be that the testing dose was too low. However, this explanation is unlikely to be true because there were 36 rapid verification tasks. Hence, the effects of adaptation via rapid assessment reported in the literature (e.g., Kalyuga & Sweller, 2004, 2005) are probably genuine adaptation effects and not just effects of the additional rapid assessment tasks.

As rapid assessment was revealed as a non-reactive assessment procedure, we used this method in a further study to identify knowledge gaps. In the case of knowledge gaps, the learners restudied the contents which they had not learned well.

Experiment 2

Incorrect answers to rapid assessment tasks might indicate that a very specific piece of knowledge is missing or that there are deficits with respect to the corresponding sub-area of the learning contents. A restudy prompt that focuses on the specific piece of knowledge tested in rapid assessment should be the most straightforward way to repair a very specific deficit, but it would not address broader deficits in the corresponding sub-area.

A more general restudy prompt should have the advantage of broader effects on the knowledge sub-area. Expecting such broader effects on learning when prompts pose more general questions is in line with the findings of Vollmeyer and Burns (2002). They took up the goal specificity effect, as revealed by cognitive load research (e.g., Sweller et al., 2011), and tested whether learners acquire more declarative knowledge when they pursue more general goals during learning in a hypermedia (multimedia) environment. Actually, it is more general goals that lead to better learning. Brunstein and Krems (2005) observed similar results in conjunction with learning from hypertext. Nevertheless, unspecific prompts or goals might also have disadvantages. If the prompts focus on the specific knowledge gaps identified by rapid assessment, these knowledge gaps might be closed more reliably than when the prompts ask for more general exploration of a knowledge area.

In this context, we assumed a differential effectiveness of specific prompts exclusively directing attention to a specific piece of knowledge and of unspecific prompts directing attention to the corresponding knowledge sub-area. Specific prompts should close knowledge gaps more effectively and they should foster the type of knowledge that is directly tapped by the rapid assessment tasks, namely the most important (central) knowledge about the mitotic process. The more general prompts should be more effective when knowledge is being fostered that extends beyond the central issues in the mitotic process (e.g., cell structure in general, or transfer tasks).

In addition, specific rather than unspecific prompts may have motivational disadvantages (Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015). If learners are given narrow limits, they may feel less autonomous. Such a loss of autonomy can reduce interest (Krapp, 2005; Niemiec & Ryan, 2009).

In summary, we addressed the following hypotheses: Specific prompts are more effective in repairing specific knowledge gaps identified by rapid assessment (*Knowledge-Repair Hypothesis)*. Specific prompts are more effective in fostering knowledge about central issues of the mitotic process (*Central-Learning-Contents Hypothesis*). Unspecific prompts are more effective in fostering knowledge about more general issues related to mitosis (*Broader-Learning Hypothesis*). Unspecific prompts are more effective in fostering situational interest (*Interest Hypothesis*).

Method

Sample and design

Forty-one university students recruited mainly from a psychology department took part in this study (age: M = 22.41, SD = 2.76). They received either study credits or 15 Euros for participation. The students were randomly assigned to two conditions: Adaptive restudy with (1) specific prompts (directing attention to the specific missing piece of knowledge; n = 20) or (2) unspecific prompts (directing attention to the knowledge area “embedding” the specific information; n = 21). The dependent measures referred to the learning outcomes in terms of repaired knowledge gaps, knowledge acquisition about the mitotic process as well as about general issues related to mitosis, and to situational interest in the learning contents.

Materials

The pretest assessing prior knowledge contained 8 items. One open item also included in the posttest asked the students to describe the process of mitosis in about 150 words. We assigned 1 point for each of the 24 aspects included in an expert answer. Two raters coded the answers for about 25% of the participants (i.e., 10 persons). We observed high interrater agreement as revealed by an ICC of .931; disagreements were resolved by discussion. Given this high agreement, the rest of the answers on this pretest and posttest item were coded by a single rater. In addition, we determined the internal consistency of this pretest measure (regarding the 24 aspects as items). We obtained a satisfying Cronbach's alpha coefficient of .79 for this measure of prior topic knowledge (on the mitotic process, i.e., the central learning contents). Seven items that had not been included in the posttest asked for other useful prior knowledge when learning about mitosis (e.g., "List five elements in the human cell."). For the open items, we again assigned 1 point for each aspect included in expert answers. Two raters coded the answers from 10 participants (i.e., about 25%). These raters also reached high agreement as indicated by an ICC of .932; disagreements were resolved by discussion. We obtained an internal consistency estimate of .73 (Cronbach's alpha) for this score of prerequisite knowledge. As both prior knowledge scores correlated with $r = .81$, $p < .001$, we computed a combined score for prior knowledge.

In the beginning of the posttest, the students worked again on all rapid assessment tasks that they had answered incorrectly during the learning phase. Thereby we could see to what extent prompted restudy closed the specific knowledge gaps. Learning outcomes on central learning contents (i.e., mitotic process) were assessed by the same item as in the pretest, that is, the students were requested to describe this process in about 150 words. We again determined a reliability estimate in terms of internal consistency (Cronbach's alpha: .83). In addition, we used 16 multiple-choice items (e.g., the learners inspected a schematic picture of the mitotic process with an error in it; they had to mark one of four presumable errors) and 5 open items (e.g., three marked structures in a realistic picture had to be named). Two raters coded the answers to the five open questions for about 25% of the participants (i.e., 10 persons). As in the case of the open pretest items, the raters referred to expert answers to score the participants' answers. We again observed high interrater agreement, as determined by an ICC of .947; disagreements were resolved by discussion. Given this high agreement, the rest of the answers were coded by a single rater. For the posttest score including the 16 multiple-choice items and the 5 open items, we determined a Cronbach's alpha of .77. Both posttest scores correlated with $r = .59$, $p < .001$, which is substantial but does not preclude separate analysis of both scores.

Situational interest was assessed by 10 items that were to be answered on a Likert scale from 1 to 7 (1: I do not at all agree; 7: I fully agree): four items referred to the value-related component of situational interest (e.g., "The topic of the learning environment is important") and six items to the emotional component (e.g., "I was bored while I worked on the learning environment"; negatively keyed item) (Schiefele & Krapp, 1996). We determined a reliability estimate of .93 (Cronbach's alpha).

Learning environment and experimental variation

We took the learning environment from Experiment 1 (i.e., same content, same number of pages, same page design, etc.). However, we made some modifications to "repair" some sub-optimal features of the text and the pictures. In addition, the participants worked on the rapid assessment tasks after smaller blocks of learning contents (e.g., each mitotic phase) in this experiment. We wanted to close the knowledge gaps more or less immediately so they would not impede further learning.

Overall, the learning environment comprised ten blocks. Each block depicted information on cell division followed by three rapid assessment tasks. These tasks presented statements, and the participants had to indicate whether the statements were right or wrong. In the case of incorrect answers, the learners were automatically directed to the corresponding learning contents for restudy in order to close their knowledge gaps.

Restudying was prompted differently between conditions. The prompts guided the learners on how to process the information on the page to which the learners had been re-directed. In the specific prompts condition, the relevant passages were highlighted by darkening the less relevant information on the page (Figure 2). Note that the darkened text passages could still be read relatively easily so that the presented information did not differ between conditions.

The prompt consisted of the request to restudy the relevant passage in order to solve the task correctly, and the task was repeated (Figure 2). In the unspecific prompts condition, the learners were asked to restudy and figure out both the direct answer to the question and to explore the broader context. For example, if the rapid assessment task "The equatorial plane is a straight plate dividing the cell during the metaphase" was answered incorrectly, the prompt "Detect what the equatorial plane is" was displayed together with the page containing the relevant information (Figure 3).

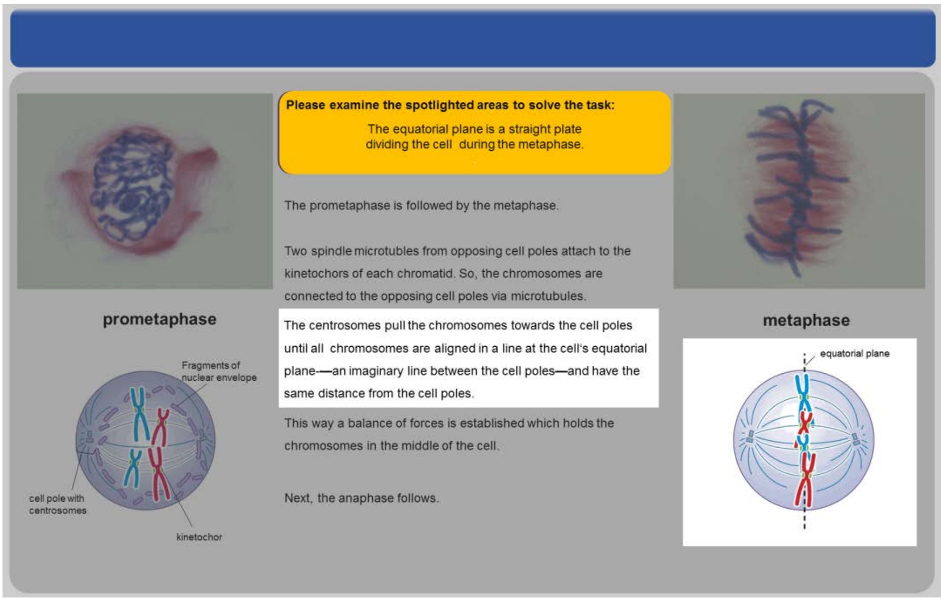

Figure 2. Screenshot with a specific prompt (Experiment 2)

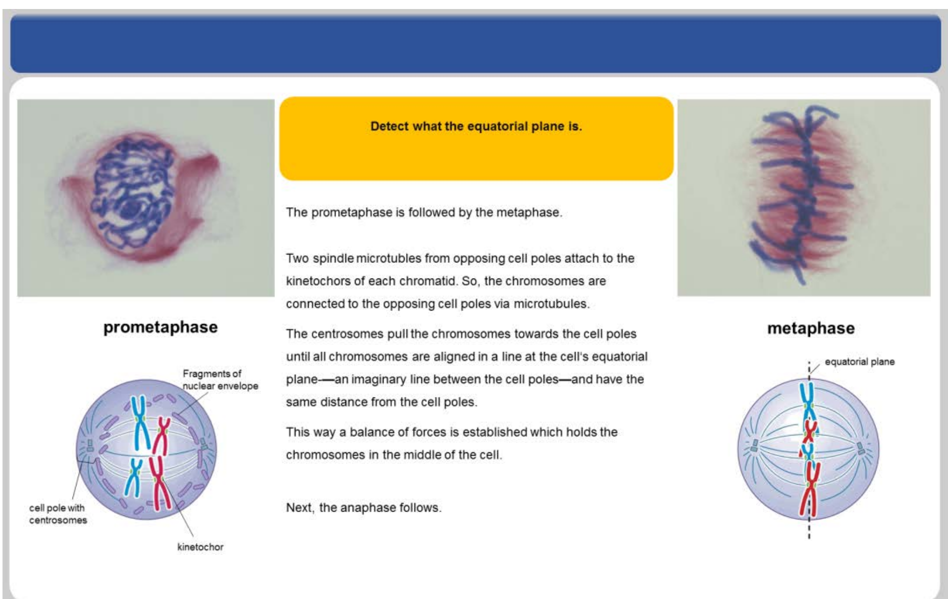

Figure 3. Screenshot with an unspecific prompt (Experiment 2)

Procedure

Participants were tested in individual sessions. The entire session was carried out on the computer. After a brief welcome, the students were familiarized with the computer system and asked to fill in a questionnaire on demographical data and to work on the pretest. Before continuing with the learning phase, participants were informed about the architecture of the learning environment, the rapid assessment tasks, and the restudy procedure.

Afterwards the students worked on the learning environment in which the experimental variation took place. After the learning environment, the students filled in the questionnaire on situational interest and worked on the posttest. At the end of these procedures, they received study credits or 15 Euros for participation.

Results

As in Experiment 1, we used an alpha level of .05, and we reported *d* as effect-size measure that was interpreted as follows: *d* = .20 as small effect, *d* = .50 as medium effect, and *d* > .80 as large effect. Table 2 shows the means and standards deviation of the central variables in both conditions. The sample size used for the statistical analyses varied slightly between 40 and 41 because of missing data.

Table 2. Means (Standard deviations in brackets) of important variables for both conditions (Experiment 2)

	Unspecific prompts (*n* = 21)	Specific prompts (*n* = 20)
Prior knowledge (z score)	0.23 (1.08)	-0.24 (0.86)
Learning time (min.)	21.10 (4.41)	21.94 (6.93)
Rapid assessment task (no. correct)	24.00 (3.36)	22.85 (3.91)
Repaired knowledge gaps (%)	82.64 (21.90)	82.07 (15.62)
Knowledge of mitotic process (no. correct)	8.81 (4.45)	5.95 (4.34)
General knowledge (no. correct)	25.88 (5.42)	22.30 (5.65)
Situational interest (scale from 1 to 7)	5.72 (0.73)	4.77 (1.58)

The two conditions did not differ significantly in their prior knowledge (overall score), *t*(39) = 1.53, *p* = .113, *d* = 0.48. However, there was a descriptive difference between conditions of almost half a standard deviation. Hence, we planned to statistically control for prior knowledge in the following analyses by ANCOVAs. However, the pretest scores did not significantly predict any of the performance measures, so that ANCOVAs made no sense. As the participants reached on average a percentage-correct score of about 14% in the pretest, they can be regarded as learners with a low level of prior knowledge.

There were no significant group differences with respect to age, semesters at university, German as mother tongue, or prior biology courses (all *p*s > .12). The participants' gender was not equally distributed across conditions (specific prompts: 19 females and 2 males; unspecific prompts: 11 females and 9 males; $chi^2(1) = 6.67$, *p* = .010). However, gender was not related to repaired knowledge gaps, knowledge about the mitotic process, and more general knowledge (all *p*s > .69). Female and male participants did not significantly differ in interest either, *t*(38) = 1.52, *p* = .137. Nor was there a difference between groups with respect to learning time (Table 2), *t*(38) = -0.46, *p* = .648.

There were 30 rapid assessment tasks. On average 23.43 (*SD* = 3.64) tasks were correctly answered. Hence, there were about 78% correct answers. The conditions did not differ in this respect (Table 2), *t*(38) = 0.99, *p* = .325. Overall, the initially incorrect answers to the rapid assessment tasks were corrected in slightly more than 82% of the cases. The two conditions did not differ significantly in how successfully the specific knowledge gaps were repaired (Table 2), *t*(38) = 0.094, *p* = .926. Both types of prompts seemed to be suited for closing knowledge gaps. Hence, we rejected the *Knowledge-Repair Hypothesis*.

With respect to learning outcomes about the mitotic process (i.e., central learning contents), we found – in contrast to our expectations – the unspecific prompts to be superior (Table 2), *t*(39) = 2.08, *p* = .044, *d* = 0.65. Hence, our *Central-Learning-Contents Hypothesis* had to be rejected. With regard to the more general learning outcomes, we found the unspecific prompts to be superior (Table 2), *t*(39) = 2.07, *p* = .045, *d* = 0.65, confirming the *Broader-Learning Hypothesis*.

Learners with unspecific prompts stated higher interest in the learning contents than those with specific prompts (Table 2), *t*(26.73) = 2.45 (test for unequal variances), *p* = .021, *d* = 0.77. We thus confirmed the *Interest Hypothesis.*

Post hoc we tested for exploratory reasons whether the superior learning outcomes of the group with unspecific prompts were mediated by enhanced situational interest. In fact, situational interest correlated substantially with both learning outcome measures (knowledge about the mitotic process: *r* = .48, *p* = .002; more general knowledge: *r*

= .54, $p < .001$). These findings point to mediation effects (Hayes, 2013). When directly testing the mediation effects by a bootstrapping procedure (number of bootstrap samples = 1000) as suggested by Hayes, we found that situational interest significantly mediated the effects of unspecific (vs. specific) restudy prompts on learning outcomes in terms of general knowledge and knowledge about the mitotic process. Unspecific prompts significantly heightened interest (b = -.95, SE = .388, t = -2.448, p = .019, LCL = -1.736, UCL = -.165), and interest was a significant predictor of general knowledge (b = 2.17, SE = .659, t = 3.306, p = .002, LCL = 0.843, UCL = 3.151) controlling for direct effects of the specificity of prompts. At the same time, the specificity of prompts (i.e., the experimental condition) lost its predictive value (b = 1.430, SE = 1.697, t = -0.843, p = .405, LCL = -4.868, UCL = 2.007) for general knowledge. A similar pattern appeared for an indirect effect of prompt specificity via interest on knowledge about the mitotic process. We found interest to be a significant predictor of knowledge about the mitotic process (b = 1.51, SE = .540, t = 2.801, p = .008, LCL = 0.419, UCL = 2.608) while the specificity of prompts lost its predictive value (b = 1.312, SE = 1.391, t = -0.944, p = .355, LCL = -4.130, UCL = 1.506). As an effect size measure for the indirect (mediation) effect, we calculated the ratio of the indirect effect of the prompt specificity (as mediated by interest) to the total effect of the prompt specificity on learning outcomes. With respect to the general-knowledge outcome variable, the indirect effect represented 59.1% of the total effect. With respect to the knowledge-about-the-mitotic-process outcome variable, the indirect effect represented 52.3 % of the total effect.

Discussion

We assumed differential effects of specific and unspecific prompts on learning outcomes. Specific prompts should be preferable to repair specific knowledge gaps and foster knowledge directly related to the mitotic process (*Knowledge-Repair Hypothesis* and *Central-Learning-Contents Hypothesis*). Unspecific prompts should be preferable in fostering knowledge about general issues related to mitosis (*Broader-Learning Hypothesis*). What we found was the general superiority of unspecific prompts, except that both types of prompts repaired the specific knowledge gaps in most cases.

In addition, our findings suggest that unspecific prompts foster situational interest. This effect is in line with self-determination theory that assumes that enhancing learners' perceived autonomy enhances motivation (e.g., Niemiec & Ryan, 2009). However, it will be up to a future study to determine whether situational interest is actually fostered by an increased level of perceived autonomy when working with unspecific prompts. In any case, our findings from the mediation analysis indicate that the positive learning effects of unspecific prompts are (partly) due to their motivational advantage.

Overall discussion

In Experiment 1, we identified no indication that rapid assessment tasks are reactive with respect to learning outcomes. This finding can be considered to be good news *and* bad news. From a practical perspective, it is bad news because it would be terrific if the diagnostic procedure underlying an adaptation procedure already had positive learning effects. From an experimental stance, it is good news because this non-reactivity makes it easier to determine pure effects of the adaptive provision of restudy prompts.

The non-reactivity of rapid assessment on learning outcomes might be surprising in light of evidence reported in the testing-effect literature (for reviews see, e.g., Rawson & Dunlosky, 2012; Roediger et al., 2011). Note, however, that we did not provide feedback in the first study. Feedback is discussed as an important ingredient when testing procedures are designed to foster learning (e.g., Rawson & Dunlosky, 2012; Roediger et al., 2011). We are well aware that a testing effect can also occur when no feedback is provided (e.g., Roediger & Karpicke, 2006). On the other hand, some testing effect studies have found that testing without feedback can even exert negative effects (i.e., consolidating incorrect knowledge: e.g., Roediger & Marsh, 2005).

Another possible cause for the present findings' divergence from the testing effect may be due to the fact that we administered an immediate posttest only. Testing effects are typically found in delayed posttests (e.g., Roediger & Karpicke, 2006). Hence, it would be sensible to use a delayed posttest in further studies.

Another limitation associated with Experiment 1 is that we did not test all possible mediation effects. For example, the rapid assessment tasks could have cancelled out learners' over-confidence or illusion of knowledge which can, in turn, foster learning outcomes (e.g., Rawson & Dunlosky, 2012). However, as there was no overall effect on learning, we tentatively assume that we did not miss an important mediational mechanism.

The positive effects of unspecific prompts on learning outcomes suggest that suggest that adaptive systems should close knowledge gaps not just by addressing the very specific knowledge gaps but by more general prompts for restudy. Such unspecific restudy has both cognitive and motivational advantages. Our findings demonstrate that the present effect of goal specificity is caused not only by cognitive factors, as suggested by prior research (e.g., Vollmeyer & Burns, 2002: learning strategies and use of resources), but by motivational factors such as interest as well.

Overall, the present experiments inform about the effects of our adaptation procedure consisting of a rapid assessment-based provision of restudy prompts. Rapid assessment does not per se have a positive effect on learning outcomes. However, such diagnostic tasks can be used to close specific knowledge gaps that have been identified (i.e., remediation rate over 80%). To exert broader effects on learning outcomes, rapid assessment should be combined with unspecific restudy prompts.

Note, however, that the performance on the learning outcome measures, even in the condition with the unspecific restudy prompts, was well below the ceiling. This sub-optimal performance might be due to the fact that our pre-determined rapid assessment tasks were not able to reveal all or at least most of the knowledge deficits that the learners had during learning. A potential remedy might be to use online data such as eye movements to collect "suspicious facts" (e.g., very little time on specific information or regressions to already-studied materials). Such indicators could be used to adaptively select rapid assessment tasks that could then verify or falsify the "suspicion," leading to the corresponding decision to present or skip a restudy prompt. Such a "double-adaptive" system (adaptive selection of rapid assessment tasks and adaptive selection of prompts) might help to further optimize learning.

Overall, we provide evidence that knowledge gaps during learning can be detected and closed by a rapid assessment-based adaptation procedure that presents (unspecific) restudy prompts. This procedure should continue to be improved in order to detect knowledge gaps during learning more sensitively. With the present two experiments, we have laid a sound foundation for future work on optimizing an adaptation procedure for expository multimedia learning environments.

References

Ackerman, P. L., & Lohman, D. F. (2006). Individual differences in cognitive functions. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 139-161). Mahwah, NJ: Erlbaum.

Ambrose, S. A., & Lovett, M. C. (2014). Prior knowledge is more important than content: Skills and beliefs also impact learning. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum.* Retrieved from http://teachpsych.org/ebooks/asle2014/index.php

Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*, 193-211.

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short-and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293-306.

Brunstein, A. & Krems, J. F. (2005). Einfluss des Bearbeitungsziels auf die Strategiewahl beim hypertextgestützten Lernen [Effects of processing goals on learning strategies for processing hypertext]. *German Journal of Educational Psychology, 19,* 39-48.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning, 28*, 557–573.

Dochy, F., de Rijdt, C., & Dyck W. (2002). Cognitive prerequisites and learning how far have we progressed since bloom? Implications for educational practice and teaching. *Active Learning in Higher Education, 3,* 265-284

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A General role of analogical encoding. *Journal of Educational Psychology*, *95*, 393-408.

Graesser, A. C., Jeon, M., & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes, 45*, 298-322.

Hayes, A. F. (2013). *An Introduction to mediation, moderation, and conditional process analysis: A Regression-based approach*. New York, NY: The Guilford Press.

Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Journal of Educational Technology and Society*, *15*, 38-52.

Kalyuga, S. (2006). Assessment of learners' organized knowledge structures in adaptive learning environments. *Applied Cognitive Psychology, 20*, 333–342.

Kalyuga, S. (2008). When less is more in cognitive diagnosis: A rapid online method for diagnosing learner task-specific expertise. *Journal of Educational Psychology, 100*, 603–612.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology, 96*, 558-568.

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development, 53*, 83–93.

Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science, 38*, 209-215.

Kalyuga S. (2012). Role of prior knowledge in learning processes. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 2886 – 2888). Berlin, Germany: Springer

Kalyuga, S., Renkl, A., & Paas, F. (2010). Facilitating flexible problem solving: A Cognitive load perspective. *Educational Psychology Review, 22*, 175–186.

Koedinger, K., & Corbett, A. (2006). Cognitive Tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61-78), New York, NY: Cambridge University Press.

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning & Instruction, 15*, 381-395.

Lee, C. H., & Kalyuga, S. (2014). Expertise reversal effect and its instructional implications. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum.* Retrieved from http://teachpsych.org/ebooks/asle2014/index.php

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, *7*, 133-144.

Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise-reversal effects in writing-to-learn. *Instructional Science, 38,* 237-258.

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning & Instruction*, *19*, 259-271.

Park, O., & Lee, J. (2003). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed., pp. 651–684). Bloomington, ID: The Association for Educational Communications and Technology.

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. E. S. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27,* 91–109.

Rawson, K. A., & Dunlosky, J. (2012). When is practice testing most effective for improving the durability and efficiency of student learning?. *Educational Psychology Review, 24*, 419–435.

Renkl, A. (2014). Towards an instructionally-oriented theory of example-based learning. *Cognitive Science, 38*, 1-37.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A Cognitive load perspective. *Educational Psychologist, 38,* 15-22.

Roediger, H. L., & Karpicke, J. D. (2006). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Roediger, H. L., & Marsh, E. J. (2005). The Positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*, 1155–1159.

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation -Advances in Research and Theory, 55*, 1-36.

Roelle, J., Berthold, K., & Renkl, A. (2014). Two instructional aids to optimise processing and learning from instructional explanations. *Instructional Science, 42*, 207–228.

Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning & Individual Differences, 8,* 141-160.

Schraw, G., Flowerday, T., & Lehman, S. (2001). Promoting situational interest in the classroom. *Educational Psychology Review*, *13,* 211-224.

Shute, V. J., & Zapata-Rivera, D. (2008). Adaptive technologies. In J. M. Spector, M. D. Merril, J. J. G. van Merriënboer, & M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 277–294). New York, NY: Taylor and Francis.

Sweller J., Ayres P. L., & Kalyuga S. (2011). *Cognitive load theory*. New York, NY: Springer. doi:10.1007/978-1-4419-8126-4

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The Contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*, 118-130.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 16-26.

Vollmeyer, R., & Burns, B. D. (2002). Goal specificity and learning with a hypermedia program. *Experimental Psychology, 49,* 98-108.