

This is the author accepted version of the article published in Learning and Instruction (including a corrigendum):

Schnaubert, L., & Bodemer, D. (2017). Prompting and visualising monitoring outcomes: guiding self-regulatory processes with confidence judgments. *Learning and Instruction, 49*, 251-262. <https://doi.org/10.1016/j.learninstruc.2017.03.004>

Licencing of this author accepted version is [CC-BY-NC-ND](#)

Title: Prompting and visualising monitoring outcomes: guiding self-regulatory processes with confidence judgments

Authors:

Lenka Schnaubert^a, Daniel Bodemer^b

^{a,b}University of Duisburg-Essen, Lotharstrasse 65, 47057 Duisburg, Germany

^alenka.schnaubert@uni-due.de (corresponding author)

^bbodemer@uni-due.de

Abstract: Sensible self-regulated study decisions are largely based on monitoring learning and using this information to control learning processes, but research has found that such processes may not be initiated automatically. To support learners, we adopted prompting and visualisation methods by asking learners to assign confidence ratings to learning tasks and visualising them during re-study, and tested the effects on metacognitive and cognitive measures in an experimental study ($N = 95$). Results show that prompting monitoring increased study efforts while visualising monitoring outcomes during learning focussed these efforts on uncertain answers. Due to low monitoring accuracy, metacognitively sensible regulation did not lead to cognitive learning gains. While the results support the idea of using visualisation techniques to implicitly guide self-regulated learning, more needs to be done to increase monitoring accuracy. Further, our study suggests that researchers should be aware of the effect that assessing confidence judgments has on subsequent learning behaviour.

Keywords: metacognitive self-regulation, response confidence, prompting monitoring, visualisation, implicit guidance

Highlights:

- learners not always monitor their learning or use the information to guide learning
- support may be realized by prompting and visualisation techniques
- prompting monitoring affected the magnitude of study behaviour
- visualising monitoring ratings focused study efforts on uncertain material
- low monitoring accuracy impeded transfer to cognitive learning outcomes

Prompting and visualising monitoring outcomes: guiding self-regulatory processes with confidence judgments

Abstract: Sensible self-regulated study decisions are largely based on monitoring learning and using this information to control learning processes, but research has found that such processes may not be initiated automatically. To support learners, we adopted prompting and visualisation methods by asking learners to assign confidence ratings to learning tasks and visualising them during re-study, and tested the effects on metacognitive and cognitive measures in an experimental study ($N = 95$). Results show that prompting monitoring increased study efforts while visualising monitoring outcomes during learning focussed these efforts on uncertain answers. Due to low monitoring accuracy, metacognitively sensible regulation did not lead to cognitive learning gains. While the results support the idea of using visualisation techniques to implicitly guide self-regulated learning, more needs to be done to increase monitoring accuracy. Further, our study suggests that researchers should be aware of the effect that assessing confidence judgments has on subsequent learning behaviour.

Keywords: metacognitive self-regulation, response confidence, prompting monitoring, visualisation, implicit guidance

1 Introduction

1.1 Metacognitive Regulation of Learning Processes

Theories on metacognitive self-regulation of learning assume a cyclic model, in which learners monitor their learning process and use this information to control learning decisions (Efklides, 2008; Nelson & Narens, 1990). According to Nelson and Narens' framework (1990), learners monitor their learning processes and outcomes (i.e., their object level) and use this information to build a dynamic, meta-level model. This model is used as information to control the learning process itself and thus in turn alters the object level. For example, learners may monitor their attempt to retrieve specific information from memory and, due to experiencing difficulties, judge the information as not learned sufficiently. Based on this information, they may decide to re-study the information altering their actual knowledge. There has been extensive research on how and how well learners monitor their learning (e.g., Maki, 1998), how they use this information to control the learning process (regulation of study, e.g., Nelson & Leonesio, 1988; Thiede, Anderson, & Therriault, 2003), and how this affects learning outcomes (e.g., Nelson, Dunlosky, Graf, & Narens, 1994; Thiede, 1999). Researchers widely assume that learners use their monitoring judgments to control studying (cf. Winne & Hadwin, 1998), and research has repeatedly produced strong evidence that learners can do so successfully (e.g., Kornell & Metcalfe, 2006; Metcalfe, 2009; Thiede, 1999). However, metacognitive self-regulation can still be very demanding and overstrain inexperienced learners (Kalyuga, 2009). Thus, in this paper, we introduce a study that investigates ways to support learning processes and outcomes by implicitly guiding self-regulation efforts based on metacognitive monitoring.

When studying, self-regulated learners have to make important decisions about their learning processes, such as what to study when, whether to continue or terminate studying or how long to study material (Nelson & Narens, 1990). According to Metcalfe and Kornell (2005), allocating study time consists of two stages: choice and perseverance. At the choice stage, learners decide which items

they need to study and the order in which to study them. Items already mastered are mostly discarded while items not yet mastered are likely candidates for study. Although different views such as the region of proximal learning framework (e.g., Metcalfe & Kornell, 2005) and discrepancy reduction views (e.g., Thiede & Dunlosky, 1999) suggest different approaches, they agree that not-yet mastered items are prioritised. At the perseverance stage, learners decide on how much time to spend on the chosen items and thus when to terminate study. All of these decisions may be based on process monitoring (Nelson et al., 1994), but can also be part of overall task goals or agendas (Dunlosky & Ariel, 2011; Dunlosky & Thiede, 2004; Thiede & Dunlosky, 1999). Even if effective agendas vary greatly depending on personal and situational factors, they all involve self-evaluation strategies to adapt study behaviour to subjective needs (Ariel, Dunlosky, & Bailey, 2009). The learner must detect such need and keep it mentally present to make study decisions accordingly. However, this might not be possible in challenging learning scenarios. Thus, there are two obstacles to regulating learning processes: the detection of a need to study and its mental presence in order to make adequate control decisions.

To detect the need to study, learners have to monitor their learning. However, such monitoring processes and the judgments that result from it (monitoring judgments) can only provide a sound basis for controlling the learning process (and thus lead to effective regulation) if they are sufficiently accurate (Dunlosky & Rawson, 2012). There are two possible cases of misjudgement: (1) overconfidence (e.g., a firm belief in the correctness of objectively incorrect information), which could lead to understudying (Dunlosky, Rawson, & Middleton, 2005) or misinformed decisions (Leclercq, 1983); and (2) underconfidence, which might yield positive results due to overlearning given unlimited resources, but might have detrimental effects if it requires that scarce resources be allocated to already mastered learning material (Dunlosky & Rawson, 2012). Regardless of the accuracy of monitoring judgments, low confidence discloses gaps in knowledge that need to be addressed to gain usable knowledge (e.g., Hunt, 2003). Therefore, learners should be likely to address uncertainties if they are aware of them.

Research has shown that actively trying to retrieve an answer from memory positively affects the accuracy of metacognitive judgments (Dunlosky et al., 2005). For example, response confidence judgments (RCJs), which require learners to evaluate their responses to learning tasks, have been shown to be more accurate in predicting actual performance than judgments made prior to retrieval attempts (e.g., Costermans, Lories, & Ansay, 1992; Maki, 1998). As discussed above, accurately monitoring performance is highly important for self-regulated learning processes and outcomes, as it influences the usefulness and the effectiveness of study decisions (Dunlosky & Rawson, 2012), such as deciding when to re-study an item or topic (Thiede, 1999). Consequently, RCJs seem to be a suitable basis for such decisions, as learners can (re-)study if they are not confident about their responses to learning tasks.

As stated, RCJs are subjective post-answer evaluations of the validity of one's own answers (i.e., subjective validity) (Leclercq, 1983) and may thus be used as a guide for further learning. While the formation of such metacognitive evaluations may be an unconscious process (Efklides, 2008), their strategic usage requires an active maintenance of the information in memory in order to compare specific evaluations (Dunlosky & Ariel, 2011) and thus conscious awareness (Efklides, 2008). This active processing consumes cognitive capacities, especially if learners must prioritise and choose between simultaneously presented materials. Item selection within simultaneously presented material activates planning activities, presumably due to automatic engagement of inter-item comparison processes necessary to make well-founded study decisions (Dunlosky & Thiede, 2004). Consequently,

metacognitive processes are related to high mental effort. While assigning cognitive resources towards sensible regulation (e.g., sensible item selection) may benefit learning by focussing attention on relevant material, it may still overstrain inexperienced learners (Kalyuga, 2009). The additional effort required by metacognitive processes may be one reason why effective regulation sometimes fails: Learners do not always actively monitor their learning (production deficiency, cf. Veenman, Kerseboom, & Imthorn, 2000; Winne, 1996) or do so only implicitly, which might result in less aware metacognitive information and thus no solid basis for control decisions. Conversely, learners might thoroughly monitor their learning but fail to use this valuable information to control learning processes, resulting in a fall-back to habitual behaviour strategies (Ariel, Al-Harthy, Was, & Dunlosky, 2011; Ariel & Dunlosky, 2012), because the metacognitive information is not readily available and hard to mentally obtain during learning. Thus, effective regulation support should address not only the lack of monitoring, but may also foster the usage of its outcome by enhancing its salience and reducing the effort of utilising this information.

1.2 Fostering Metacognitive Self-Regulation

Metacognitive self-regulation may fail if learners are not able or not willing to monitor their learning appropriately. There are various methods to overcome availability deficiencies of monitoring, such as strategy training (e.g., Nietfeld, Cao, & Osborne, 2006), which have been successfully used to improve deficient monitoring skills. Production deficiencies, in contrast, happen when available behaviour is not executed, for example due to distraction (cf. Veenman et al., 2000). Here, direct instruction may be used more cautiously to allow for individual regulation (cf. *assistance dilemma*, Koedinger & Alevan, 2007) and instructional methods can be limited to an activation of favourable processes, e.g., by prompting. Prompting has repeatedly been found to be an effective means to support self-regulated learning (Bannert & Reimann, 2012; Wirth, 2009). Metacognitive prompts merely stimulate recall or execution of skills and thus do not teach new information (Bannert, 2009), but they do put emphasis on specific processes or concepts. A mandatory judgment on monitoring outcomes, for example, asks the learner to monitor their cognitive processes explicitly and to externalise the outcome by rating it on a given scale. Following these prompts thus triggers monitoring and additionally makes the outcome more salient. Recent research has shown that monitoring judgments, i.e. judgments of learning, are highly reactive, affecting for example study time allocation (Mitchum, Kelley, & Fox, 2016) or memory (Soderstrom, Clark, Halamish, & Bjork, 2015). While judgments of learning are assumed to foster an active memory search, which may act as rehearsal in case of successful recall, RCJs do not serve this function since they refer to already retrieved answers. Thus, it remains unclear whether assessing RCJs influences self-regulated study processes.

Whilst monitoring processes have been prompted successfully in the past, promoting their usage to guide study decisions seems more difficult. As we discussed earlier, adequate control strategies even though available (e.g., choosing appropriate items to study) might fail if the task exceeds the mental capacities of the learners. Computational systems offer the possibility to permanently take study decisions off the learners' hands (e.g., Kornell & Metcalfe, 2006; Nelson et al., 1994), but this digresses far from the idea of self-regulated and autonomous learners. Thus, support strategies are needed that relieve the cognitive system while tacitly guiding the learners' self-regulation attempts. One strategy, borrowed from group awareness research, is the salient visualisation of knowledge-related information to support learners in structuring their common learning processes (Janssen & Bodemer, 2013); this includes visualisations of metacognitive judgments (Dehler, Bodemer, Buder, & Hesse, 2011). By providing salient, easily comparable visualisations of (lacks of) knowledge, such tools may

guide learning while still enabling a self-directed approach (Bodemer, 2011). Previous work conducted in group awareness research allowed for comparisons of group members' knowledge or knowledge gaps via visualisations, but this approach may also be suitable for individual learning. Making monitoring outcomes like RCJs externally available by saliently visualising them to ease comparison processes between studied items might also foster metacognitive self-regulation. In metacognitive self-regulation the focus is on inter-item rather than inter-individual comparison processes, because the former are necessary for choosing one item over the other (cf. Dunlosky & Thiede, 2004). Such visualisations may act as visual markers, signalling material that needs further attention, without the need of constant mental availability of the information. Similar to group awareness tools, the information provided should be easy to understand and interpret to prevent distraction (Bodemer & Dehler, 2011). Note, however, that such visualisation techniques cannot be addressed entirely separately from prompting, as these visualisations require assessing monitoring outcomes, which in turn requires explicitly asking learners to monitor themselves.

1.3 Research Question and Hypotheses

Our aim is to investigate ways to support self-regulatory learning processes. More specifically, we want to know: Can we facilitate metacognitive self-regulated learning by prompting monitoring processes and by providing external representations of these mental constructs? We will focus on response confidence judgments (RCJs) attached to responses to specific learning tasks as the target concept. As specified in section 1.1, RCJs serve an important function in monitoring memory and thus learning outcomes and consequently this information is a valuable indicator for the necessity to (re-)study material. We are interested in different dependent measures focussing on learning processes as well as outcomes.

1.3.1 Learning Processes

In general, we are interested in the learners' study behaviour and how this relates to their monitoring outcomes. By prompting learners to monitor their learning outcomes in the form of RCJs, learners may be more aware of their uncertainties and perceived lacks of knowledge. Thus, they may feel the need to increase study efforts by studying more information. However, if they are not supported in using their RCJs to focus their efforts towards specific items (e.g., by visualisation), these efforts may be unfocussed such as searching for information to validate most of their responses. As argued above, visualisation, in contrast, should support learners to utilise their monitoring outcomes and thus conduct a more focussed approach to studying by making it an easily accessible learning strategy to study mainly material that the learners are unsure about (cf. section 1.1). However, since especially the effects of prompting on such quantitative aspects of study behaviour like the amount of material chosen to study are largely unknown, we abstain from formulating a unidirectional hypothesis, but cautiously assume that prompting and visualisation may affect the quantity of study behaviour (hypothesis 1).

Apart from quantitative aspects, we are also interested in qualitative aspects of study decisions. Based on the research mentioned above, we assume that learners in general use their monitoring outcomes (i.e., confidence) to control their learning, since research has shown that learners choose items to re-study based on monitoring outcomes (e.g., Thiede et al., 2003). In accordance with the argumentation in section 1.2, we assume that being asked to provide explicit RCJs to responses to learning tasks prompts monitoring processes. Additionally, the externalisation of the RCJs should increase the salience of the judgments. Both prompting and externalisation processes should thus lead to a better utilisation of these judgments to make study decisions. Having monitoring judgments readily available

during learning through visualisations may enhance this effect by making own monitoring outcomes even more salient and facilitating inter-item comparison processes by relieving working memory. Thus, we assume that learners who are prompted to monitor their memory use their RCJs to make study decisions (metacognitive regulation of study). Being particularly supported in utilising the RCJs by external visualisations should enhance this effect (hypothesis 2). Assuming that monitoring outcomes to some extent reflect the actual state of learning (cf. section 1.1), an approach focussing on items solved with low confidence (uncertain items) should also lead primarily to the selection of objectively needed information, i.e., information regarding incorrect responses (hypothesis 3). However, this effect should be smaller than the one specified in hypothesis 2, because monitoring accuracy is a potentially limiting factor.

Additionally, we assume that learners supported by visualisations not only use their ratings to decide what to study, but also when to study it (prioritising, cf. Metcalfe & Kornell, 2005). We assume that they favour information regarding uncertain responses and will not only primarily study such information, but also prioritise it before information regarding certain responses. Again, while learners who are prompted to monitor their memory should attend to uncertain responses first, learners who are additionally supported by visualisations should be more consistent in the usage of their monitoring outcomes to prioritise due to its higher salience (hypothesis 4).

Since learners not only base study decisions on monitoring outcomes, but also allocate more time to items judged as not yet mastered (cf. Metcalfe & Kornell, 2005; Nelson & Leonesio, 1988), we assume that learners allocate more study time to information regarding responses judged as uncertain than to those judged as certain. Again, this effect should be greater if learners have their monitoring outcomes externally available during learning (hypothesis 5). A more focussed approach should also have another effect on study time allocation: if item selection is systematically based on the subjective or objective need to study, quick re-checks (selecting unneeded information) may be avoided and thus study times of selected information should increase as the support that learners receive increases (hypothesis 6).

However, prompting might also have some unwanted side effects. Such additional tasks may interrupt the students' learning processes (Dempsey & Driscoll, 1996) since they require switching between task-related activities (e.g., comprehending the question, retrieving information from memory) and meta-level activities (e.g., actively evaluating the answers during task completion and transforming the experience of one's own confidence to a given scale). Apart from interruption, monitoring and externalising the outcomes are activities that may strain the cognitive system. The visualisations (externalised outcomes) also need resources to be processed, but since they are external representations of metacognitive concepts, they additionally have the potential to relieve the cognitive system by focussing attention and externally providing information relevant for metacognitive regulation (cf. section 1.2). Thus, there are indications to argue for more as well as less strain on the cognitive system and consequently we abstain from formulating a unidirectional hypothesis, but merely assume that prompting monitoring and visualising the outcome affect the mental effort of learners (hypothesis 7).

1.3.2 Learning Outcomes

We assume that prompting and, even more so, visualising RCJs enables focusing of study effort. Since research has repeatedly found links between regulation of study and performance (e.g., Nelson et al., 1994; Thiede, 1999; cf. section 1.1), we assume that by altering learning processes we will foster

learning outcomes. We therefore assume knowledge gain during learning to be greater the more a learner is supported (hypothesis 8). Since we assume that prompting and, even more so, visualisation helps learners to focus on and therefore clear up uncertainties, we expect higher post-learning confidence levels for supported learners, especially learners who are able to strategically work through uncertain items due to support by visualisations (hypothesis 9).

Finally, re-studying material might not only help learners to correct faulty or uncertain knowledge, it might also impact monitoring accuracy. Learners aware of their monitoring judgments may use re-study trials to explicitly adjust faulty monitoring decisions. Consequently, we assume that learners who are prompted to externalise their monitoring outcomes will improve their monitoring accuracy and judge their knowledge more accurately. Again, we expect that learners who have their monitoring ratings readily available during learning outperform merely prompted learners (hypothesis 10).

2 Method

2.1 Sample, Design and Procedure

To test our hypotheses, we conducted an experimental study with $N = 96$ university students. In the course of the study, one participant had to be excluded due to a server error, which left us with $N = 95$ participants in the final sample. They were all university students predominantly enrolled in a Bachelors or Masters course on Applied Cognitive and Media Science (24 males, 71 females) with a mean age of 22.09 ($SD = 2.81$). Topic specific interest regarding the topics addressed in the learning material (blood sugar regulation and diabetes mellitus) measured on a scale from 0 (no interest) to 5 (high interest) was at a medium level throughout the sample (blood sugar regulation: $M = 2.57$, $SD = 0.13$; diabetes mellitus: $M = 2.58$, $SD = 0.14$); Self-assessed prior knowledge measured on a scale from 0 (low knowledge) to 5 (high knowledge) was rather low (blood sugar regulation: $M = 0.87$, $SD = 0.09$; diabetes mellitus: $M = 0.86$, $SD = 0.10$) (cf. section 2.2 for more information on the scales used). All experiments were conducted in our research lab; instructions were given via computer. Participants were rewarded either 12 Euros or course credit for research participation. Before starting the experiment, participants were randomly assigned to one of three experimental conditions. Participants in the prompting+visualisation (referred to as “visualisation” hereafter) condition assigned RCJs to learning tasks and had this information displayed during re-study. Participants in the prompting condition assigned the RCJs, but were not given this information during re-study. Participants in the control condition did not assign RCJs and consequently were not provided with such information during re-study. There were no significant differences in topic-specific interest (blood sugar regulation: $F(2,92) = 0.15$, $p = .858$, $\eta^2 < .01$; diabetes mellitus: $F(2,92) = 0.10$, $p = .909$, $\eta^2 < .01$) or self-assessed prior knowledge (blood-sugar regulation: $F(2,92) = 1.17$, $p = .315$, $\eta^2 = .03$; diabetes mellitus: $F(2,92) = 0.20$, $p = .816$, $\eta^2 < .01$) between the groups.

After participants were briefed and had given consent to participate in the study, they were provided with the experimental material on a computer screen. They were asked to give demographic information and rated their prior knowledge and interest regarding the topics addressed in the learning material (blood sugar regulation and diabetes mellitus). Then they received textual material about these topics (learning phase one, LP1) and answered learning tasks (with or without RCJs) (t1). Afterwards, they had the opportunity to (re-)study material regarding specific tasks (with or without a visual representation of their RCJs); the tasks were presented in a simultaneous format. During this second learning phase (LP2), they were able to change their answers (and depending on condition their

RCJs) (t2). Finally, they answered the learning tasks again from scratch (all with RCJs) (t3) and took a knowledge test on the learned material. After each phase, learners answered an item assessing self-reported mental effort. Figure 1 represents the overall procedure and highlights the points at which the independent variable was manipulated.

2.2 Material

The demographics questionnaire collected information on age, sex, university course and semester as well as on two variables to control for pre-test differences on blood sugar regulation and diabetes mellitus. These variables were topic-specific interest (“I think the topic diabetes mellitus [blood sugar regulation] is...”) assessed on a 6-point scale from “not interesting at all” (0) to “very interesting” (5) and prior knowledge (“My knowledge about diabetes mellitus [blood sugar regulation] is ...”) assessed on a 6-point scale from “very low” (0) to “very high” (5).

A three-page expository text (1425 words) was used to provide each student with background information on the topics (LP1). 20 learning tasks were designed to capture important aspects of these topics (Note that the term “learning task” is used to stress that – from the learners’ perspective – they are used within the learning process. However, they may still be used to assess the learners’ knowledge about the material). While some tasks directly referred to information given in the text, others referred to information not previously provided. The learning tasks each consisted of a statement that the learners were asked to verify or falsify (true-false) and were given in an array format. Translated sample items with confidence ratings are depicted in Figure 2. Depending upon point in time (t1, t2 or t3) and condition, learners were or were not additionally asked to judge their confidence in their answer. The answers were spatially coded (top – true, bottom – false) and confidence judgements were colour coded (filled green – sure, hatched green – unsure), cf. Figure 2. In t1 and t3, learning tasks were initially blank, in t2 the learners were provided with their own answers from t1 (and depending on condition, with or without respective confidence ratings).

In LP2, learners were able to request additional information on each learning task individually by clicking a button placed next to each task. The provided information was presented by an overlay window and was either taken from the initial text or consisted of new information (cf. Figure 3).

The knowledge test consisted of 25 single choice items with four alternative answers each, designed to test more profound knowledge of the information given. In contrast to the learning tasks, they were not used within the learning process and thus were used to measure if possible learning gains may be transferred to unstudied tasks. To assess individual confidence, 6-point confidence scales (“How sure are you that your answer is correct?”), ranging from “not sure at all” (0) to “absolutely sure” (5) were added. In the current sample, item difficulty was normally distributed ($S-W = .96$, $df = 25$, $p = .413$) ranging from .08 to .98 ($M = .51$, $SD = .19$) across items.

Learning material and test items were specifically developed for this study in a recursive process, testing the material on small groups of students at a time. The basis for the material was basic literature on blood sugar regulation and diabetes mellitus including common misconceptions. The development was supported by an educationalist on medicine.

Reported mental effort was assessed after LP2 and after answering the first set of learning tasks (t1) to assess differences imposed by the treatment. We used one item asking the learners how demanding the learning phase had been on a 7-point Likert scale, adapted from the mental load scale from Tindall-

Ford and colleagues (Tindall-Ford, Chandler, & Sweller, 1997), ranging from “not demanding at all” (0) to “very demanding” (6).

2.3 Independent Variables

The main independent variable was the level of metacognitive support the learners received. One group received no such support, whereas two groups were asked to provide a binary confidence rating along with each learning task in t1 in order to prompt metacognitive monitoring processes. One of these groups additionally had these ratings visualised in LP2 to support their usage for the control of study behaviour. This procedure left us with three groups: no support (control), mere prompting of monitoring processes by asking for confidence ratings (prompting), and additional visualisation of said confidence ratings during learning (visualisation). Since the support factor was progressively staggered, Helmert contrasts were used to separate the impact of general support, prompting and visualisation. Additionally, some measures were taken repeatedly, e.g., performance and confidence regarding the learning tasks were measured at two points within the experiment (t1, t3). This left us with a two factorial design with one within- and one between-subjects factor.

2.4 Dependent Variables

To measure how metacognitive support affects cognitive learning gain (hypothesis 8), we measured performance in the learning tasks (t1, t3) and the knowledge test (sum of correctly solved items). Additionally, we were interested in the impact that support had on metacognitive measures. We assessed confidence levels by counting the confidently solved items in the learning tasks (t1 for prompting and visualisation condition, t3) and by computing mean confidence ratings for the knowledge test (independent of correctness of answers) (hypothesis 9). Further, to assess how well learners monitor themselves (hypothesis 10), we computed relative accuracy measures in the form of individual within-subject phi- or Goodman-Kruskal’s gamma-coefficients between performance and confidence ratings in the learning tasks (t1 for prompting and visualisation condition, t3) or the knowledge test (cf. Schraw, Kuch, & Gutierrez, 2013). High positive coefficients indicate good monitoring accuracy since learners tend to be confident when they are also correct and not confident when they are incorrect, while negative indices imply the opposite.

Addressing the quantity of study behaviour (hypothesis 1), we assessed how many learning tasks the students requested information for by counting non-recurring information requests. Additionally, we were interested in qualitative aspects of study behaviour, i.e. how learners in the prompting and in the visualisation conditions used their confidence ratings to make study decisions (metacognitive regulation of study, hypothesis 2). Therefore we computed within-subject phi-coefficients between initial confidence (t1) and information requests (LP2), a method frequently used to assess metacognitive regulation of study (e.g., Thiede, 1999). High positive indices indicate that learners mainly assess information about uncertain items (good metacognitive regulation), while negative indices indicate the opposite. Coefficients near zero indicate no differentiation between certain and uncertain items. To see if learners made objectively useful study decisions, we also computed phi-coefficients between performance and information requests to compare between all groups (objective quality, hypothesis 3). We recoded the data to ensure that high coefficients again mean useful study decisions (requesting mainly information on incorrect answers).

To capture the influence of confidence on the order of study requests (hypothesis 4), we used an algorithm designed to measure the time-wise prioritisation of non-confident or confident responses with regard to information requests per learner (Schraubert & Bodemer, 2016). By computing

individual mean-rank differences between confidence levels, we ensured that the number of appearance of each level did not affect the index. The index ranges from +10 (all uncertain items are considered before certain items) to -10 (all certain items are considered first) and has a theoretical mean of 0 (no prioritisation).

With regard to study time allocation, we assessed study durations per requested information (hypothesis 6). We also tested how study time allocation depended on initial confidence for the two conditions that provided confidence ratings prior to learning, by measuring mean study durations per confidence level, including only items for which information was requested (hypothesis 5).

To assess if the support changed the mental effort needed by the learners (hypothesis 7), we compared reported mental effort between the conditions at two points in time: After initial task completion, we compared the prompted conditions with the non-prompted condition, and after LP2 we compared all three conditions.

There were no significant correlations between performance at the beginning of the study (learning tasks t1) and dependent process variables (e.g., number of information requests, metacognitive regulation of study, objective quality of study decisions, etc.) or monitoring accuracy. Thus, we assume this influence on the results to be negligible.

3 Results

To answer our research questions, we conducted several analyses according to distribution assumptions on the dependent variables. If not specified otherwise, results of Shapiro-Wilk tests did not contradict the normality assumption and we therefore used parametric analyses. We also conducted planned contrasts (Helmert) to take into account the staggered arrangement of the metacognitive support. We conducted two-tailed analyses to allow for opposing effects, level of statistical significance was set at $\alpha = .05$.

3.1 Learning Processes

In the following, we discuss the results concerning learning processes. We focus on the quantity of study behaviour first (3.1.1), followed by two sections on item selection (choice) namely quality of study decisions (3.1.2) and order of processing (3.1.3), and one section on the actual allocation of study time (perseverance, 3.1.4). Finally, we report on the effects on reported mental effort (3.1.5).

3.1.1 Quantity of Study Behaviour (hypothesis 1)

First, we looked at the quantity of study behaviour (number of information requests). Descriptive statistics are provided in table 1. A Welch-Test showed no significant difference between the conditions regarding the quantity of study behaviour ($F(2, 58.87) = 2.97, p = .059; \eta^2 = .06$). However, Helmert contrasts revealed a significant difference between the non-prompted and both prompted conditions ($t(63.29) = 2.45, p = .017, d = 0.53$), but not between the two prompted conditions ($t(49.44) = -0.60, p = .549, d = 0.15$).

Table 1: Descriptive statistics on number of information requests per condition

condition	number of information requests		
	<i>N</i>	<i>M</i>	<i>SD</i>
control	32	8.13	4.63
prompting	32	11.00	6.03
visualisation	31	10.26	3.43
overall	95	9.79	4.93

3.1.2 Quality of Study Decisions (Regulation of Study; hypotheses 2 & 3)

In a second step, we were interested in how learners used confidence ratings to make their study decisions (hypothesis 2). Descriptive analyses of the phi-coefficients between initial confidence and study requests show a median of .17 (*IQR* = .46) for the prompting and .74 (*IQR* = .47) for the visualisation condition (the control condition did not provide confidence ratings at t1 and thus had to be excluded from analyses regarding hypothesis 2). Due to violations of the normality assumption, we conducted a Mann-Whitney-U-test, which revealed a significant difference in study regulation between the two groups ($U = 62.00$, $Z = -.5739$, $p < .001$, $r = .07$). A Wilcoxon signed rank test confirmed a significant deviation from zero for the prompting ($Z = -2.443$, $p = .015$, $r = .43$) as well as for the visualisation condition ($Z = -4.870$, $p < .001$, $r = .87$), meaning that both groups used their confidence ratings to make study decision, though to a different extent. In contrast, analyses on objective quality of study decisions (hypothesis 3; cf. table 2) showed no inter-group-differences ($F(2, 89) = 0.41$, $p = .667$, $\eta^2 = .01$) as well as no significant difference from zero for the whole sample ($N = 92$; $t(91) = -0.50$, $p = .616$, $d = 0.05$).

Table 2: Descriptive statistics on objective quality of study decisions per condition

condition	objective quality		
	<i>N</i>	<i>M</i>	<i>SD</i>
control	32	-.005	.222
prompting	29	-.002	.239
visualisation	31	.042	.231
overall	92	.012	.229

3.1.3 Order of Processing (hypothesis 4)

We then conducted the sequence analyses to assess whether learners attended to uncertain or certain items first – again only with the groups providing confidence ratings prior to learning. Wilcoxon signed rank test showed no significant deviation from zero for the mean rank differences of the prompting condition ($Z = 1.117$, $p = .264$, $r = .20$), but for the visualisation condition ($Z = 4.880$, $p < .001$, $r = .88$). A Mann-Whitney-U-Test (conducted due to violations of the normality assumption) revealed a significant difference between the groups ($U = 923.50$, $Z = 5.890$, $p < .001$, $r = .74$) with the visualisation condition having a significantly higher mean rank difference in favour of uncertain items ($Mdn = 7.50$, $IQR = 4.55$) than the prompting condition ($Mdn = -0.21$, $IQR = 4.37$).

3.1.4 Study Time Allocation (hypotheses 5 & 6)

While the results described in sections 3.1.2 and 3.1.3 are concerned with item choices, we additionally were interested in how learners allocate study time to those chosen items (hypothesis 6) and if they further differentiate between confidently and not confidently solved items (hypothesis 5). With regard to hypothesis 6, we found that mean study durations per item did differ between the three groups

($F(2, 92) = 5.57, p = .005, \eta^2 = .108$). Helmert contrasts revealed that there was a difference between the prompted and not prompted conditions ($t(92) = 2.56, p = .012, d = 0.56$) with prompted learners spending more time per additional information, as well as between the prompted conditions ($t(92) = 2.16, p = .033, d = 0.54$), again with more support leading to longer study durations (cf. table 3). Due to the significance of the second contrast, we also contrasted the control and the prompting only condition to extract the prompting effect. A t -test for independent samples revealed no significant difference between these conditions ($t(62) = 1.19, p = .239, d = 0.30$).

Table 3: Descriptive statistics on mean study durations per requested information per condition

condition	study duration per request (sec)		
	<i>N</i>	<i>M</i>	<i>SD</i>
control	32	23.84	9.59
prompting	32	27.45	14.20
visualisation	31	34.35	13.76
overall	95	28.49	13.28

With regard to hypothesis 5, we tested how study-time allocation depended on initial confidence for the two conditions that provided confidence ratings prior to the second learning phase (t1). A two-factorial ANOVA with repeated measures on one factor was administered to test the effects of initial confidence and condition on the mean study duration per selected additional information (cf. figure 4). The results show a significant main effect of confidence ($F(1, 46) = 9.05, p = .004, \eta_p^2 = .16$), no main effect of condition ($F(1, 46) = 0.10, p = .753, \eta_p^2 < .01$), and no significant interaction ($F(1, 46) = 3.62, p = .063, \eta_p^2 = .07$). N differs from other calculations due to specific study patterns of fourteen participants (who did not select any confidently solved item for re-study) and the elimination of an extreme value in the visualisation group.

3.1.5 Reported Mental Effort (hypothesis 7)

We assessed reported mental effort with one item. Non-parametrical Mann-Whitney-U-tests revealed no difference in the load imposed by the first set of learning tasks between the prompted and non-prompted conditions ($U = 959.00, Z = -.393, p = .694, r = .04$), with both groups reporting a medium load ($Mdn = 3.00, IQR = 2.00; M_{non-prompted} = 3.03, SD_{non-prompted} = 1.60; M_{prompted} = 2.89, SD_{prompted} = 1.43$). Comparing the overall load imposed on the learners by learning phase two also showed no difference among the three conditions (Kruskal-Wallis-test: $H = .866, df = 2, p = .649$) and also not between visualisation and non-visualisation conditions ($U = 1097.00, Z = .852, p = .394, r = .09$). The visualisation condition reported a load of $Mdn = 3.00 (IQR = 2.00, M = 2.42, SD = 1.39)$, and the others of $Mdn = 2.00 (IQR = 2.00; M_{prompting} = 2.22, SD_{prompting} = 1.48; M_{control} = 2.31, SD_{control} = 1.47)$.

3.2 Learning outcomes

In the following sections, we present the results regarding our research questions on knowledge (3.2.1), confidence levels (3.2.2) and monitoring accuracy (3.3.3).

3.2.1 Task- and Test-Performance (hypothesis 8)

The mean number of correctly solved items in the knowledge test did not differ among the conditions ($F(2, 92) = 0.03, p = .972, \eta^2 = .001$). However, a two-factorial ANOVA on the number of correctly solved items in the learning tasks revealed a significant main effect of time, with all groups performing

significantly better after than before learning phase two ($F(1, 92) = 61.20, p < .001, \eta_p^2 = .40$). There was no significant main effect of condition ($F(2, 92) = 2.16, p = .121, \eta_p^2 = .05$) or an interaction ($F(2, 92) = 0.11, p = .894, \eta_p^2 < .01$). The descriptive statistics are available in table 4.

Table 4: Descriptive statistics on knowledge (test performance) per condition

condition	N	performance knowledge test (max. = 25)		performance learning tasks pre (max. = 20)		performance learning tasks post (max. = 20)	
		M	SD	M	SD	M	SD
control	32	12.72	3.14	10.59	1.83	12.38	1.81
prompting	32	12.78	3.51	10.34	2.24	12.41	2.33
visualisation	31	12.90	2.71	11.23	2.06	13.13	1.93
overall	95	12.80	3.11	10.72	2.06	10.72	2.06

3.2.2 Confidence Level (hypothesis 9)

Mean confidence levels in the knowledge tests were roughly in the middle of the scale for all three conditions ($M_{control} = 2.79, SD_{control} = 0.67, M_{prompting} = 2.73, SD_{prompting} = 0.71, M_{visualisation} = 2.76, SD_{visualisation} = 0.68$) and there was no significant difference among the groups ($F(2, 92) = 0.07, p = .934, \eta^2 < .01$). However, there was a significant difference among the groups in the learning tasks post learning phase two ($F(2, 92) = 11.43, p < .001, \eta^2 = .20$) (cf. t3 in figure 5). Helmert contrast revealed a significant effect between both supported conditions and the not supported condition, with learners in the supported conditions being more confident ($t(92) = 4.10, p < .001, d = 0.89$). Additionally, there was a significant difference between the two supported conditions ($t(92) = 2.49, p = .014, d = 0.63$), with learners in the visualisation condition being more confident than those in the prompted only condition. Due to the significance of the second contrast, we contrasted the control and the prompting condition to extract the prompting effect. A t -test for independent samples revealed a significant difference between these conditions ($t(62) = 2.19, p = .032, d = 0.58$), with learners in the prompted condition being more confident than those in the control condition. Two-factorial analyses between the two prompted conditions (prompting only and visualisation) revealed a highly significant effect of time ($F(1, 61) = 194.98, p < .001, \eta_p^2 = .76$) with the participants becoming more certain from pre to post LP2. Additionally, it showed a significant interaction between time and condition ($F(1, 61) = 9.14, p = .004, \eta_p^2 = .13$) with learners in the visualisation condition gaining more confidence than learners in the prompting only condition. There was no main effect of condition ($F(1, 61) = 1.27, p = .265, \eta_p^2 = .02$) (cf. figure 5).

3.2.3 Monitoring Accuracy (hypothesis 10)

Analyses on monitoring accuracy showed that with respect to the learning tasks, phi-coefficients were generally low (cf. table 5) and did not differ between the three groups post re-study ($F(2, 78) = 0.62, p = .540, \eta^2 = .02$). Further, a two-factorial repeated-measures ANOVA for both prompted conditions revealed neither a significant effect of time ($F(1, 47) = 0.85, p = .360, \eta_p^2 = .02$) nor of condition ($F(1, 47) = 0.28, p = .598, \eta_p^2 = .01$), nor an interaction ($F(1, 47) = 0.02, p = .898, \eta_p^2 < .01$). As for the knowledge test, a one-way ANOVA showed no differences in the gamma-coefficients between the three groups ($F(2, 92) = 0.03, p = .974, \eta^2 < .01$). Descriptive statistics for the accuracy measures are provided in table 5.

Table 5: Descriptive statistics on monitoring accuracy (within-subject correlations between certainty and performance) per condition

condition	N	Gamma knowledge test		Phi learning tasks pre		Phi learning tasks post		
		M	SD	M	SD	N ⁽¹⁾	M	SD
control	32	.37***	.25	--	--	32	.18***	.24
prompting	32	.39***	.23	.10 ⁰	.30	28	.14**	.22
visualisation	31	.38***	.24	.02 ⁰	.25	21	.11*	.21
overall	95	.38***	.24	--	--	81	.15***	.22
all prompt	63	.39***	.23	.06 ⁰	.28	49	.13***	.21

⁽¹⁾ N varies due to stability on the confidence dimension in the learning tasks post (certain in all items)
Significance of the means' deviation from zero: ⁰p > .05; *p < .05; **p < .01; ***p < .001

3.2.4 Interrelation between Dependent Variables

Since most of our dependent variables are assumed to be interconnected, we modelled these interactions via a moderated mediation (cf. figure 6 for the statistical model, coefficients are based on z-scores). Results show a significant overall model for explaining objective quality of study regulation through metacognitive regulation moderated by monitoring accuracy ($F(3, 56) = 34.50$, $p < .001$, $R^2 = .52$) as well as a significant overall model explaining performance gain in the learning tasks from pre- to post-learning ($F(2, 57) = 4.66$, $p = .013$, $R^2 = .15$) through objective quality of regulation. Regression coefficients confirm that the effect of metacognitive regulation on learning gain is mediated through objective regulation and thus indirect only. However, monitoring accuracy moderates the relationship between metacognitive and objective regulation (first stage moderated mediation). To describe the mediation effect, we calculated the impact of metacognitive regulation on objective regulation at three different levels of monitoring accuracy. The analyses showed that values one standard deviation below the mean in monitoring accuracy resulted in a slightly negative effect of metacognitive regulation on objective regulation. Mean values in monitoring accuracy resulted in near to no effect, whereas values one standard deviation above the mean resulted in a clearly positive effect of metacognitive regulation on objective quality of regulation.

4 Discussion and Conclusion

The aim of our study was to experimentally research two ways of guiding self-regulated learning: by prompting monitoring judgments (asking for binary, item-based confidence ratings) and visualising the resulting ratings during learning. With our experiment, we showed that visualisations, especially, are suitable to foster the utilisation of monitoring judgments and thus may be used to support metacognitive regulation of study.

As expected, prompting primarily affected quantitative aspects of study behaviour and visualisation primarily affected its direction, leading to a more focussed approach. Learners adapted their behaviour to their monitoring outcomes, especially if provided with visualisation, but failed to study objectively sensible items (i.e., items they were unable to solve correctly). Accordingly, learners cleared up more uncertainties if supported by prompting and visualising techniques, but test performance was not affected. This lack of effect on objective values can be explained by a moderated mediation model. The low monitoring accuracy we found in our study hampered the subjectively sensible regulation attempts (studying primarily uncertain items) from leading to objectively sensible decisions (studying incorrectly answered items) and thus to better learning outcomes. Since there is no direct relation between metacognitive study regulation and learning outcomes, changes in study regulation depend

on monitoring accuracy to take effect. If monitoring and performance are not related, this detaches the meta-level from the object level, leading –in our case– to completely sensible behaviour and behavioural outcomes from a subject-centred perspective, but not from an outside perspective. Our results suggest that learners were either not able or not willing to precisely monitor their learning, which may have been partly due to the fact that the learning material included common misconceptions on diabetes mellitus, making accurate monitoring even harder. If learners lack the metacognitive skills to effectively use tacit regulatory support, there might be more need to explicitly support the learners' monitoring processes. Apart from the possibility of a lack of skill, prompting may also disrupt the learning processes (Bannert & Reimann, 2012; Dempsey & Driscoll, 1996). If perceived as a distraction, learners might limit the effort put into the monitoring judgments, limiting their usefulness in the long run. The learners' perception of usefulness might be a moderating factor in the usage of support provided and should be considered explicitly in further research.

While monitoring accuracy certainly was a limiting factor in the usefulness of the provided support, the actual extent of the problem cannot be fully captured by the data assessed. Monitoring accuracy measures with regard to the learning tasks might have been tainted by the 50% chance of guessing correctly hampering the validity of monitoring indices. By using binary items to assess and display monitoring judgments, we took an uncommon decision in metacognitive research. Usually, metacognitive ratings are primarily used to assess metacognitive processes or outcomes, while in our study we fed them back to the learners as implicit guidance. Therefore, we used binary ratings (tasks as well as confidence ratings) to support ease of understanding and interpretation by limiting the complexity of the design. However, such measures also limit the conclusiveness of the results. The visualisation does suggest to decide between need and no need for further attention, but also ignores the possibility of more fine-grained usage of metacognitive ratings, for example to plan and prioritise items according to pre-set goals (Ariel et al., 2009). There are various strategies for how to approach learning material based on discrepancy reduction (Thiede & Dunlosky, 1999) or a region of proximal learning approach (Metcalfe & Kornell, 2005), both with different implications for learning. We took the decision to leave sufficient time to access additional material (up to 20 minutes) with the option to end the process earlier. This procedure is realistic for self-regulated learning scenarios, as there are time constraints yet learners basically decide how long to study. However, narrower time constraints may alter strategic approaches. For example, strategies can shift during learning if time is running out (shift-to-easier-material, Dunlosky & Thiede, 2004). Son and Sethi (2006) argue that the nature of the learning curve as well as time constraints impact optimal learning strategies. Again, due to the binarity of our confidence ratings, we cannot differentiate those strategies to analyse study decisions in more detail, but we need to be aware that visualisations simplify complex concepts and focus attention towards specific aspects of metacognition (in our study, the simplification was maximised for salience and comparability). Thus, depending on how the information is pre-processed and displayed, visualisations may be more suggestive of one strategy than the other. It is possible that this design hampers more advanced processing and it might be useful to scale up the design in a further study, trying to find an optimal level weighing grain-size and complexity. More research is needed to investigate the effects of how gathering and visualising information affects the way the information is perceived and used (Buder, 2011) and how this can be used to best support learning processes. Thus, necessary next steps to take are developing scales that best represent learners' metacognitive status to gather valid, reliable and useful information and combining them with ways to visualise this information for most efficient utilisation that matches the needs of the learners. Additionally, further studies should include measures to analyse the nature of the learning curve as well as the strategic

approach in combination with different visualisation methods to guide learners towards effective and meaningful study decision. Simultaneously, it would be an asset to know whether learners actually perceive such metacognitive visualisations as helpful and disencumbering.

The results of our study support findings of studies conducted with judgments of learnings that have shown that judgments might not be explicitly generated automatically, but only be constructed in response to the trigger question (cf. Mitchum et al., 2016; Soderstrom et al., 2015). This prompting function has been shown to alter learning processes for judgments of learning and our study supports this notion for RCJs. This raises the relevant question of the external validity of metacognitive research building on self-report judgments. While literature on metacognition has addressed shortcomings of subjective judgments – for example, Winne (2010) described a variety of self-report shortcomings with regard to self-regulated learning – empirical studies often fail to explicitly acknowledge that asking for monitoring judgments does prompt learners to evaluate their learning. If habitual learning behaviour is targeted by the research conducted, the reactivity of the design is hard to argue with. Thus, understanding the prompting effect of monitoring judgments is essential in order to assess and quantify its impact on metacognitive research.

Self-report is not only problematic because of its possible reactivity. The validity of self-report measures is questionable and our study relied heavily on self-report judgments. Metacognitive judgments for example require learners to assess their metacognitive status and transform it to a given scale. While this process may flaw the outcome to some extent, it still targets the to-be-assessed concept directly (metacognitive judgments aim directly at assessing the learners' subjective view on cognition, not at assessing cognition itself; cf., Nelson & Narens, 1990), making self-report less problematic. This is different for mental effort, as the target concept (mental effort) does not directly equal the assessed variable (subjective perception of effort). More direct measures of mental effort like dual-task methodology are not appropriate for testing instructional methods, since they divert resources away from the primary task (e.g., Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Future studies should explicitly target the effects on mental effort by using physiological measures or dual-task methodology to more accurately assess how the additional monitoring activity and the visualisation of monitoring outcomes affects mental effort. Further, a subjective one-item sensor can only be a rough indicator for actual effort needed. Although one-item sensors for perceived mental effort have been shown to be reliable and sensitive measures (Paas, Merriënboer, & Adam, 1994), we cannot conclude with certainty that our treatments did not affect the mental strain put on learners. Additionally, a one-item measurement prevents distinctions from being made between cognitive resources dedicated to learning and to interfering activities. With more possibilities to differentiate between different sources of mental effort, studies could focus on interventions targeting processes which may be useful for some (e.g., inexperienced) learners, but prove detrimental for others (e.g., experienced learners). Thus, the effort involved increases for very different reasons (cf., Kalyuga, Ayres, Chandler, & Sweller, 2003). For example, for learners who are spontaneously and continuously monitoring their learning, being additionally asked to transform this experience on a given scale might cause detrimental redundancies, while for inexperienced learners it might trigger beneficial monitoring processes. Differential effects are not limited to prompting: On the one hand, visualisations might be helpful for some learners, relieving their cognitive system of efforts to re-construct this information on the fly to use it to direct their learning (resources which might then be freed up for learning processes). On the other hand, visualisations might just add load for learners who do abstain from using this particular information to direct their learning or prefer or even need the information in a different form (different grain-size, more-dimensional etc.). Thus, effects might differ for learners

according to their cognitive abilities. For example, differences in working memory capacity may affect how heavily learners rely on externalisations to relieve their working memory. Some learners supported by externalisations may profit mainly from the working memory relief, while others may profit more from the guiding effects of externalisations. It was beyond the scope of our study to extract the effects of guiding and cognitive relief, but further studies may focus directly on the specific mechanisms involved and take cognitive resources into account.

Another methodological limitation of our study is that the sample consisted of university students only. While we acknowledge that university students may not validly represent the whole population of learners, we have no reason to assume this sample to differ greatly from other university students (with an exception of students of medicine, which were excluded due to the medical topic involved). Thus, our results may not apply to non-university students and should be replicated with other populations, especially a sample with a different educational background. Since our intervention is designed to support learners building on their own competencies, this may be even harder for learners with less metacognitive skills. We can assume that university students may – overall – possess higher metacognitive skills than the general population due to their experience in (successful) self-regulated study, so the support may have greater effects on this sample. On the other hand, it may also interfere with already developed scripts that learners have established and might thus work better with less experienced learners. Our study was not designed to answer these questions and further studies should replicate these findings with other populations and integrate variables to explain possible differences (e.g., metacognitive skills, working memory capacity, intelligence or prior knowledge).

While we investigated self-regulated learning in a very individualistic environment, modern learning is not done in isolation, but highly affected by others via social learning scenarios (e.g., using social media as a source). Mixing methods from collaborative research with metacognitive research is a step towards merging those fields. Recently, (self-)regulatory processes have been integrated in models of collaborative learning (Järvelä & Hadwin, 2013) and visualisations may be used to support such scenarios (Miller & Hadwin, 2015). Providing information on learners' metacognitive evaluations may not only inform the learner her/himself, but may also trigger essential co- and shared regulation processes. In turn, other learners may be a valid source of information supporting learners in identifying gaps in knowledge or misconceptions and thus supporting monitoring. Explicitly integrating social context into metacognitive self-regulation research and metacognitive research into collaborative learning is an obvious conclusion and should increasingly be addressed in research.

The overall goal of this study was to find ways to support learners in their own self-regulation efforts. This research is especially relevant when we consider how learning has changed during the last decades. In contrast to very explicit and “enforced” methods to externally structure learning processes, the focus has shifted towards empowering learners and supporting their self-regulated learning processes. Thus, enabling learners to make relevant and sensible decisions during self-regulated learning themselves is vital and our results suggest that prompting and visualising monitoring judgments may at least support some of the required processes. However, prompting monitoring and visualising the outcomes may not only be applied by teachers as a method to train learners to incorporate such strategies into their learning processes. Tools to prompt and visualise monitoring may easily be included in digital textbooks or web-based learning scenarios, enabling students on a larger scale to take control over their learning processes without falling back to habitual, non-reflective behavioural patterns due to limited cognitive capacities or convenience. While much research is done to improve control-based monitoring, research on finding ways to foster the utilisation of monitoring

to control learning (monitoring-based control) is still scarce. Thus, the results of our study suggest that assessing and visualising monitoring judgments may be one avenue to explore further and – in combination with interventions to improve monitoring accuracy – may tacitly guide students' self-regulated learning. Such an approach has the potential to enable learners to remain agents of their own learning (cf. Hacker, Dunlosky, & Graesser, 2009) – with adequate support to make informed study decisions.

5 References

- Ariel, R., Al-Harthy, I. S., Was, C. A., & Dunlosky, J. (2011). Habitual reading biases in the allocation of study time. *Psychonomic Bulletin & Review*, 18(5), 1015–1021. <https://doi.org/10.3758/s13423-011-0128-3>
- Ariel, R., & Dunlosky, J. (2012). When do learners shift from habitual to agenda-based processes when selecting items for study? *Memory & Cognition*, 41(3), 416–428. <https://doi.org/10.3758/s13421-012-0267-4>
- Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, 138(3), 432–447. <https://doi.org/10.1037/a0015928>
- Bannert, M. (2009). Promoting self-regulated learning through prompts: A discussion. *Zeitschrift für Pädagogische Psychologie*, 23(2), 139–145. <https://doi.org/10.1024/1010-0652.23.2.139>
- Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1), 193–211. <https://doi.org/10.1007/s11251-011-9167-4>
- Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, 27(3), 1079–1086. <https://doi.org/10.1016/j.chb.2010.05.016>
- Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, 27(3), 1043–1045. <https://doi.org/10.1016/j.chb.2010.07.014>
- Buder, J. (2011). Group awareness tools for learning: Current and future directions. *Computers in Human Behavior*, 27(3), 1114–1117. <https://doi.org/10.1016/j.chb.2010.07.012>
- Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(1), 142–150. <https://doi.org/10.1037/0278-7393.18.1.142>
- Dehler, J., Bodemer, D., Buder, J., & Hesse, W. F. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, 27(3), 1068–1078. <https://doi.org/10.1016/j.chb.2010.05.018>
- Dempsey, J. V., & Driscoll, M. P. (1996). Error and feedback: Relation between content analysis and confidence of response. *Psychological Reports*, 78(3c), 1079–1089. <https://doi.org/10.2466/pr0.1996.78.3c.1079>
- Dunlosky, J., & Ariel, R. (2011). Self-regulated learning and the allocation of study time. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 54, pp. 103–140). Amsterdam, NL: Academic Press.
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Dunlosky, J., Rawson, K. A., & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language*, 52(4), 551–565. <https://doi.org/10.1016/j.jml.2005.01.011>
- Dunlosky, J., & Thiede, K. W. (2004). Causes and constraints of the shift-to-easier-materials effect in the control of study. *Memory & Cognition*, 32(5), 779–788. <https://doi.org/10.3758/BF03195868>

- Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*(4), 277–287. <https://doi.org/10.1027/1016-9040.13.4.277>
- Hacker, D. J., Dunlosky, J., & Graesser, A. C. (2009). A growing sense of “agency.” In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 1–4). New York, NY: Routledge.
- Hunt, D. P. (2003). The concept of knowledge and how to measure it. *Journal of Intellectual Capital, 4*(1), 100–113. <https://doi.org/10.1108/14691930310455414>
- Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist, 48*(1), 40–55. <https://doi.org/10.1080/00461520.2012.749153>
- Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*(1), 25–39. <https://doi.org/10.1080/00461520.2012.748006>
- Kalyuga, S. (2009). Instructional designs for the development of transferable knowledge and skills: A cognitive load perspective. *Computers in Human Behavior, 25*(2), 332–338. <https://doi.org/10.1016/j.chb.2008.12.019>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4
- Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239–264. <https://doi.org/10.1007/s10648-007-9049-0>
- Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(3), 609–622. <https://doi.org/10.1037/0278-7393.32.3.609>
- Leclercq, D. (1983). Confidence marking: Its use in testing. *Evaluation in Education, 6*(2), 161–287. [https://doi.org/10.1016/0191-765X\(82\)90011-8](https://doi.org/10.1016/0191-765X(82)90011-8)
- Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117–144). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159–163. <https://doi.org/10.1111/j.1467-8721.2009.01628.x>
- Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52*(4), 463–477. <https://doi.org/10.1016/j.jml.2004.12.001>
- Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior, 52*, 573–588. <https://doi.org/10.1016/j.chb.2015.01.050>
- Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General, 145*(2), 200–219. <https://doi.org/10.1037/a0039923>
- Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science, 5*(4), 207–213. <https://doi.org/10.1111/j.1467-9280.1994.tb00502.x>
- Nelson, T. O., & Leonesio, J. R. (1988). Allocation of self-paced study time and the “labor-in-vain effect.” *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(4), 676–686. <https://doi.org/10.1037/0278-7393.14.4.676>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of learning & motivation* (Vol. 26, pp. 125–173). New York, NY: Academic Press.
- Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*(2), 159–179. <https://doi.org/10.1007/s10409-006-9595-6>

- Paas, F., Merriënboer, J. J. G. van, & Adam, J. J. (1994). Measurement of Cognitive Load in Instructional Research. *Perceptual and Motor Skills*, 79(1), 419–430. <https://doi.org/10.2466/pms.1994.79.1.419>
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance Cognitive Load Theory. *Educational Psychologist*, 38(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8
- Schnaubert, L., & Bodemer, D. (2016). How socio-cognitive information affects individual study decisions. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016* (pp. 274–281). Singapore, SG: International Society of the Learning Sciences.
- Schraw, G., Kuch, F., & Gutierrez, A. P. (2013). Measure for measure: Calibrating ten commonly used calibration scores. *Learning and Instruction*, 24, 48–57. <https://doi.org/10.1016/j.learninstruc.2012.08.007>
- Soderstrom, N. C., Clark, C. T., Halamish, V., & Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(2), 553–558. <https://doi.org/10.1037/a0038388>
- Son, L. K., & Sethi, R. (2006). Metacognitive control and optimal learning. *Cognitive Science*, 30(4), 759–774. https://doi.org/10.1207/s15516709cog0000_74
- Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, 6(4), 662–667. <https://doi.org/10.3758/BF03212976>
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95(1), 66–73. <https://doi.org/10.1037/0022-0663.95.1.66>
- Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 1024–1037. <https://doi.org/10.1037/0278-7393.25.4.1024>
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3(4), 257–287. <https://doi.org/10.1037/1076-898X.3.4.257>
- Veenman, M. V. J., Kerseboom, L., & Imthorn, C. (2000). Test anxiety and metacognitive skillfulness: Availability versus production deficiencies. *Anxiety, Stress, & Coping*, 13(4), 391–412. <https://doi.org/10.1080/10615800008248343>
- Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, 8(4), 327–353. [https://doi.org/10.1016/S1041-6080\(96\)90022-9](https://doi.org/10.1016/S1041-6080(96)90022-9)
- Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist*, 45(4), 267–276. <https://doi.org/10.1080/00461520.2010.517150>
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum.
- Wirth, J. (Guest Ed.). (2009). Promoting self-regulated learning through prompts [Special issue]. *Zeitschrift für Pädagogische Psychologie*, 23(2), 91–145.

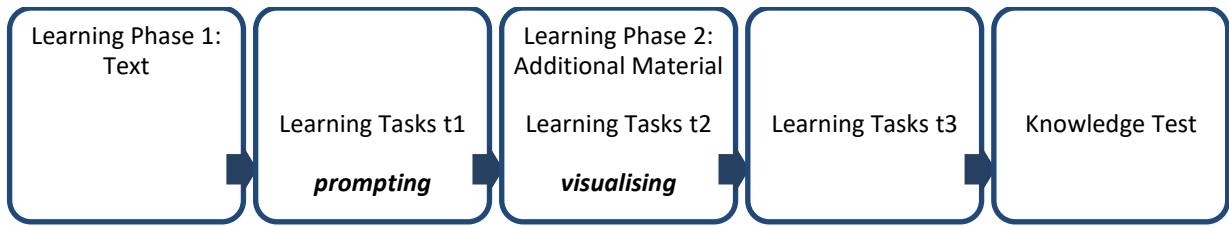


Figure 1. Overall procedure.

Type 1 diabetics produce more insulin than metabolically healthy people.	<input checked="" type="radio"/> true <input type="radio"/> false	<input checked="" type="checkbox"/>
The consumption of alcohol can cause hyperglycemia within diabetics.	<input type="radio"/> true <input checked="" type="radio"/> false	<input type="checkbox"/>

How sure are you, that your answer is correct?

sure

unsure

Figure 2. Learning tasks with confidence ratings.

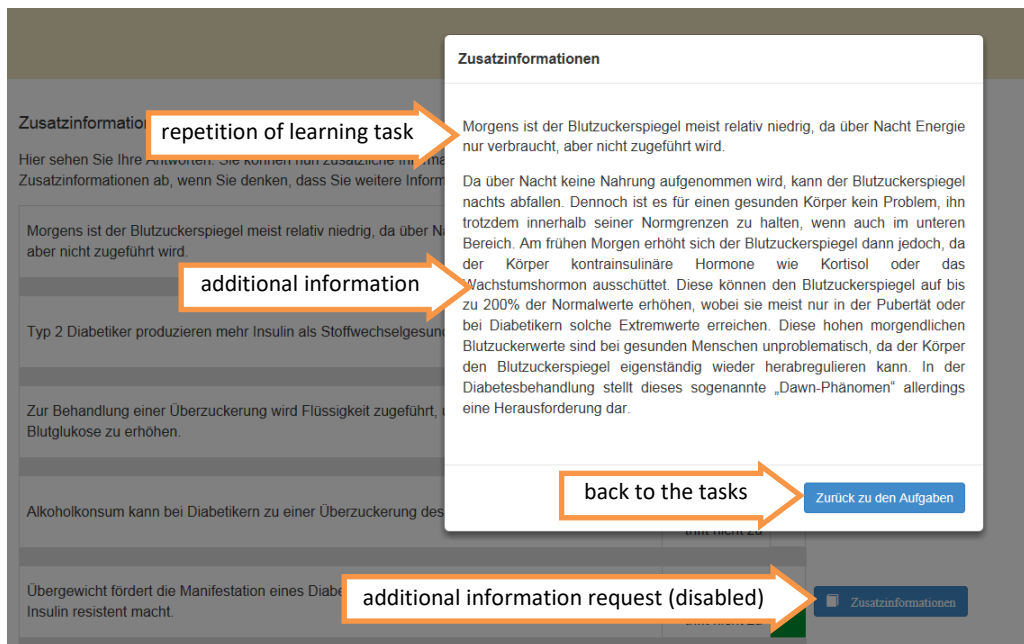


Figure 3. Additional information.

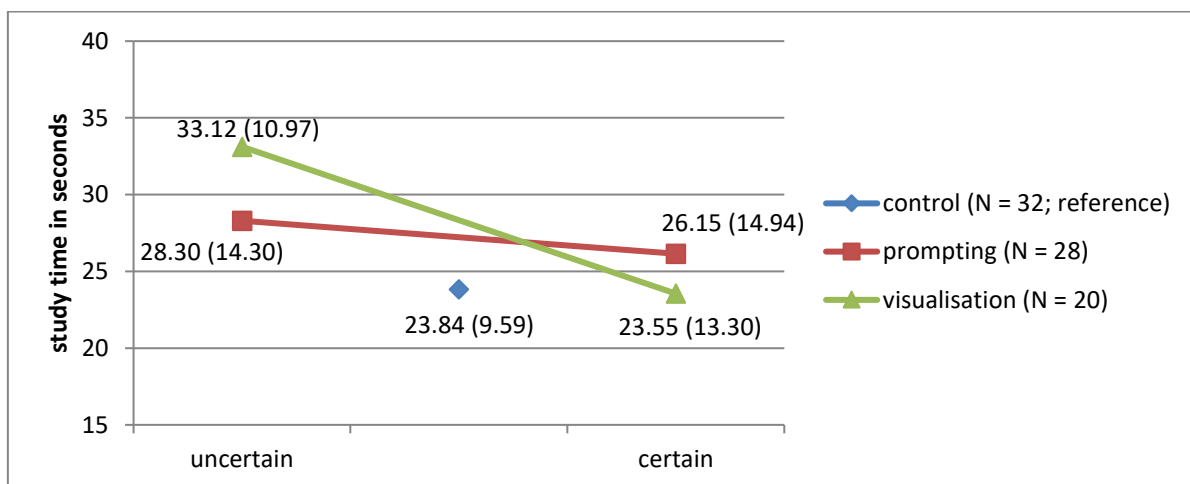


Figure 4. Means (standard deviations) of study time in seconds per chosen certain or uncertain item.

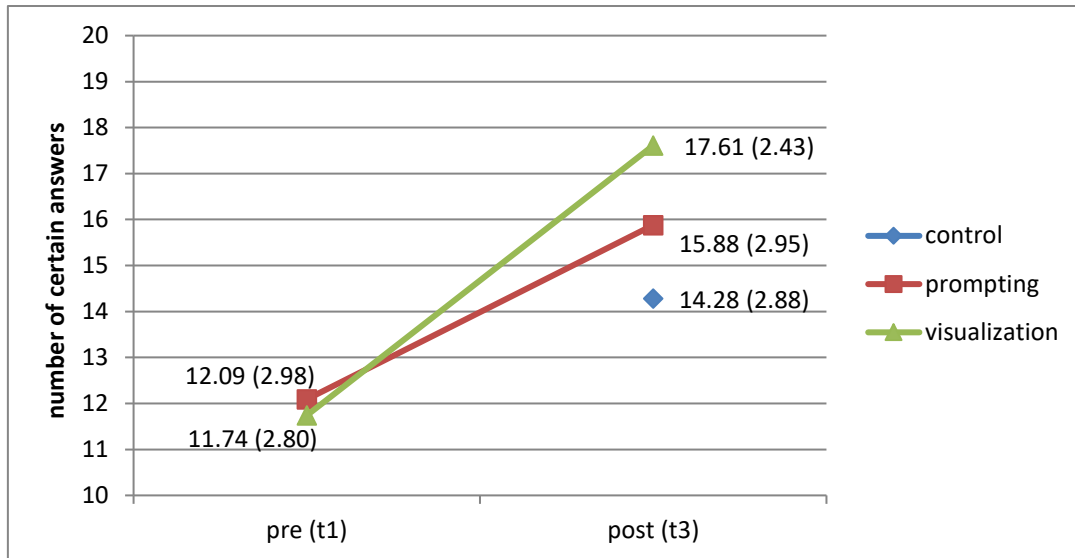


Figure 5. Means (standard deviations) of number of certain answers to learning tasks pre and post re-study.

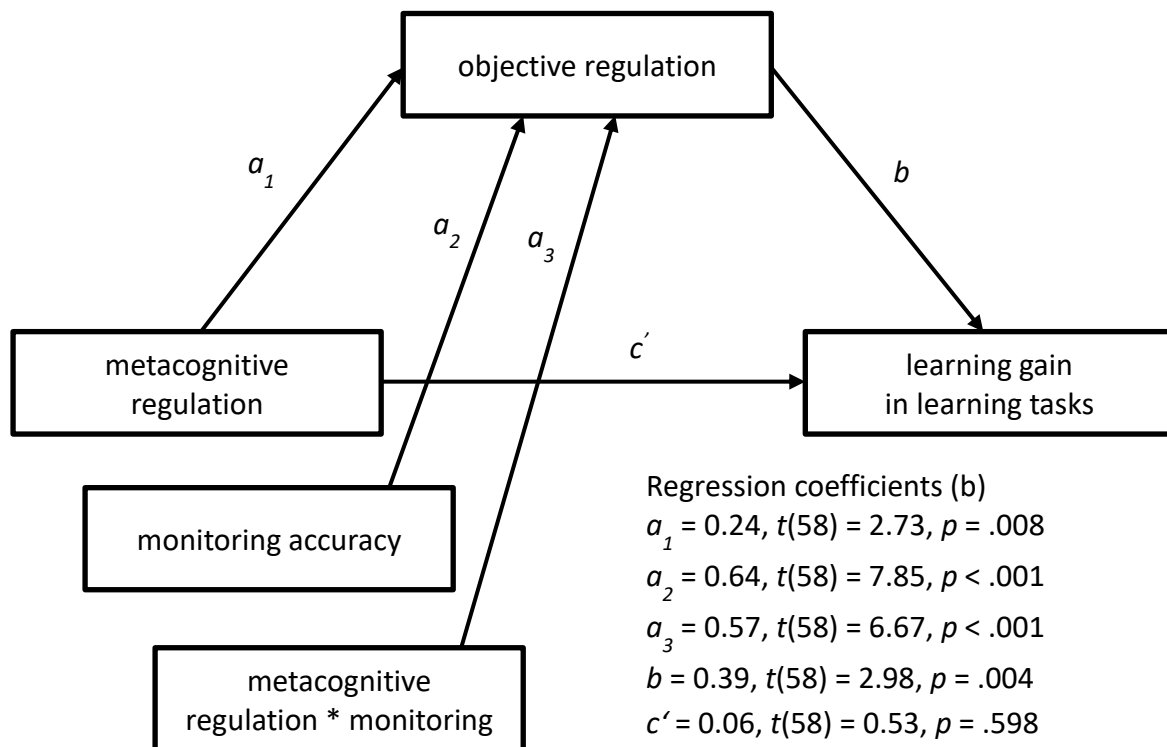


Figure 6. Statistical model of the moderated mediation: metacognitive regulation explaining learning gain via objective regulation moderated by monitoring accuracy.

Corrigendum:

Corrigendum to [Prompting and visualising monitoring outcomes: Guiding self-regulatory processes with confidence judgments].

[Journal: Learning and Instruction –Volume 59 and page range of the article]:251-262

Author details and affiliation:

Lenka Schnaubert, Daniel Bodemer*

University of Duisburg-Essen, Lotharstrasse 65, 47057 Duisburg, Germany

Corrigendum Text:

In section 3.1.2., the inferential data regarding hypothesis 2 contains an error: the data in the paper underestimates the Z-score and the effect size r by a decimal place. It should read: “Due to violations of the normality assumption, we conducted a Mann-Whitney-U-test, which revealed a significant difference in study regulation between the two groups ($U = 62.00$, $Z = -5.739$, $p < 0.001$, $r = 0.74$).”

Furthermore, in section 3.2.1. table 4 states descriptive data of task- and test-performance. However, the information designating the variables contains an error: the headings “performance knowledge test” and “performance learning tasks pre” have been interchanged. The corrected version of the table is shown below:

Table 4

Descriptive statistics on knowledge (test performance) per condition.

condition	N	performance knowledge test (max. = 25)		performance learning tasks pre (max. = 20)		performance learning tasks post (max. = 20)	
		M	SD	M	SD	M	SD
control	32	12.72	3.14	10.59	1.83	12.38	1.81
prompting	32	12.78	3.51	10.34	2.24	12.41	2.33
visualisation	31	12.90	2.71	11.23	2.06	13.13	1.93
overall	95	12.80	3.11	10.72	2.06	10.72	2.06