Structure, relationship, and determinants of monitoring strategies and judgment accuracy. An integrated model and evidence from two studies☆

Marion Händel[a,b,1,*], Markus Dresel[b]

[a] Department of Psychology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
[b] Department of Psychology, University of Augsburg, Germany

A B S T R A C T

Metacognitive monitoring plays a central role in models focusing on either monitoring strategies (models of self-regulated learning) or monitoring judgments and their accuracy (models of metacognition). Although monitoring strategies and monitoring judgments are both concerned with monitoring one's learning progress, they have been analyzed independently so far. To combine these two research perspectives, we propose an integrated model. Two studies empirically tested the factorial structure of metacognitive monitoring and investigated the influence of metacognitive knowledge and motivational components. The studies focused on global (Study 1, $N = 396$) and situation-specific (Study 2, $N = 225$) metacognitive monitoring as well as quantitative and qualitative aspects of metacognitive monitoring (both studies). Metacognitive monitoring was characterized by three separate but correlated factors: quantity of monitoring strategy use, quality of monitoring strategy use, and judgment accuracy. Furthermore, common (metacognitive knowledge, expectancy of success) and specific (attainment value) determinants of metacognitive monitoring were identified.

Imagine a learner who monitors her understanding and learning progress. For example, she tries to explain the content she is learning to a peer in her own words in order to test her knowledge, or she estimates that a certain number of tasks have been solved incorrectly in a sample exam. While the first behavior represents the use of a monitoring strategy, the second behavior represents the generation of a monitoring judgment. If the learner recognizes difficulties in understanding, she should, for example, deepen her knowledge, elaborate on the content, do further literature research, or change her strategy. As a consequence, monitoring is supposed to influence future learning and—due to the use of more appropriate cognitive learning strategies like elaboration or retrieval—also learning success (Kostons, van Gog, & Paas, 2012; Metcalfe & Finn, 2008).

Metacognitive monitoring is central to models of self-regulated learning (Winne & Hadwin, 2008; Zimmerman, 2008) as well as models of metacognition (Flavell, 1979; Nelson & Narens, 1990). For an overview of different models, see Panadero (2017) who reviewed and compared the six most prominent models of self-regulated learning. Previous research on metacognitive monitoring was predominantly either part of primarily strategy-oriented research (i.e. research that mainly focusses on the use of self-regulated learning strategies; see Winne & Hadwin, 2008) or primarily judgment-oriented research (i.e. research that mainly focuses on monitoring judgments as the procedural part of metacognition; Koriat, 1997). Monitoring strategies are—next to planning and regulation—one type of metacognitive or so-called self-regulatory strategies (Pintrich, 1999). Monitoring strategies inherit a strategic component like engaging in a specific task with the aim to monitor one's own cognitive activities and actual behavior. Monitoring judgments display a self-evaluative decision about whether a learner thinks that he or she has understood a specific concept or not. Table 1 provides an overview of different aspects of metacognitive monitoring. Despite their strong conceptual overlap, the research literature and theoretical models are largely unconnected, and monitoring strategies and monitoring judgments have thus far mostly been analyzed independently. Only a handful of studies have investigated monitoring strategy use and monitoring judgments simultaneously (Bong, 1997; Gidalevich & Kramarski, 2017; Saraç & Karakelle, 2017; Sperling, Howard, Staley, & DuBois, 2004). These studies provide only limited evidence on the interplay of monitoring strategies and monitoring judgments as some of them did not focus on monitoring strategies

☆ The authors declare that they have no conflict of interest.

* Corresponding author at: Ansbach University of Applied Sciences, Faculty of Media, Residenzstraße 8, 91522 Ansbach, Germany.
E-mail address: marion.haendel@hs-ansbach.de (M. Händel).
1 Marion Händel is now at Ansbach University of Applied Sciences, Germany.

Table 1
Overview of the different aspects of metacognitive monitoring.

Term	Description
Metacognitive monitoring	Umbrella term for monitoring learning or monitoring understanding. Regarded as a "key form of self-regulation", "involves observing and tracking one's own performance and outcomes" (Zimmerman, 1998, p. 78). Distinguished from metacognitive control (Nelson & Narens, 1990; Winne & Hadwin, 1998), which modifies learning.
Monitoring strategies	One type of self-regulated learning strategy. Strategies as self-testing or reviewing with the aim to self-evaluate learning behavior and understanding. Distinguished from planning and regulation strategies in models of self-regulated learning (Pintrich, 1999).
Quantity of monitoring strategy use	Frequency of used strategies to monitor one's individual learning and understanding.
Quality of monitoring strategy use	Quality of used strategies to monitor one's individual learning and understanding. Refers to target orientation, precision, and control of monitoring strategies (Engelschalk, Steuer, & Dresel, 2017).
Monitoring judgments	Judgments regarding one's individual performance in a task. Can be global (e.g., judgment regarding a whole test) or item-specific (one specific judgment reading a task or a knowledge item in a test). Can be made before (prediction) or after (postdiction) learning or testing.
Judgment accuracy	Indicator for the quality of monitoring judgments. Can be measured as absolute (e.g., absolute accuracy, bias) or relative (e.g., gamma) accuracy score.

specifically but rather an umbrella of several metacognitive strategies like planning, self-evaluation, or regulation (Craig, Hale, Grainger, & Stewart, 2020; Griffin, Wiley, & Salas, 2013). Others of the aforementioned studies did not investigate the two concepts from an integrated perspective. Finally, those studies used quite diverse populations ranging from primary to higher education students. This might explain the heterogenous results. Hence, there is no elaborated understanding on how monitoring strategies and monitoring judgments are related, whether they constitute distinct aspects of monitoring, or whether they are just "a rose by any other name", as titled by Lajoie (2008, p. 469). To advance the theoretical understanding and to provide further evidence on metacognitive monitoring as part of theoretical models of metacognition and self-regulated learning (Dinsmore, Alexander, & Loughlin, 2008; Flavell, 1981; Kaplan, 2008), the current work pursues an integrated approach. This means to integrate the strategy-oriented perspective and the judgment-oriented perspective of metacognitive monitoring by simultaneously modelling monitoring strategies and judgment accuracy as well as their interplay. In addition, potential determinants of both perspectives are investigated.

1. Metacognitive monitoring

Metacognitive monitoring is "one's on-line awareness of comprehension and task performance" (Schraw & Moshman, 1995, p. 355). Metacognitive monitoring can occur in different phases in the learning process (i.e. forethought phase, performance phase, and self-reflection phase; Zimmerman, 1998), can refer to different aspects of learning (i.e., monitoring strategy use, monitoring understanding, monitoring goals; Greene & Azevedo, 2007), and can refer to learning processes (e.g., monitoring understanding during reading) as well as to learning products (e.g., monitoring the consistency of an finished written essay; Pintrich, 2004; Sobocinski et al., 2020). In addition, metacognitive monitoring is understood as a situation-specific and context-dependent process (Boekaerts, 1999; Dresel et al., 2015; Koriat, 2019; McCardle & Hadwin, 2015). For example, during and after learning, students can monitor their learning progress by elaborating on how well they have reached their set goals (Winne & Hadwin, 1998). In order to regulate learning adequately, this "[monitoring] process by which the metalevel tracks the accuracy of object level-performance" (Koriat, 2019, p. 2) should be as accurate as possible.

In the following, we elaborate theoretically on monitoring strategies and monitoring judgments and their interplay (Table 1 provides definitions for the core aspects of metacognitive monitoring). Specifically, we differentiate quantity and quality aspects of metacognitive monitoring, discuss the relationship of monitoring strategies and judgment accuracy, as well as common (relevant for both, monitoring strategies and judgment accuracy) and specific (relevant for just one of the two aspects) determinants of metacognitive monitoring. A common determinant describes a determinant that is relevant for monitoring strategies as well as judgment accuracy. In contrast, a specific determinant is only relevant for either monitoring strategies or judgment accuracy.

Based on these theoretical considerations, we propose an integrated theoretical model on these factors of metacognitive monitoring and their interplay (Fig. 1).

1.1. *Monitoring strategies: types, quantity, and quality of use*

Monitoring strategies can concern a broad set of objects (goals, strategy use, learning approach, learning progress, motivation and emotion, or learning results). In this work, we focus on monitoring one's understanding and learning performance, that is, cognitive monitoring according to Winne and Hadwin (1998). In this sense, monitoring strategies are cognitive or behavioral strategies with the aim to evaluate learning behavior and understanding (Flavell, 1979; Winne & Hadwin, 2008). Typical monitoring strategies are self-testing/self-questioning as well as reviewing during the learning process with the aim to prove understanding or to determine concepts that are not understood well (Lan, 2005; Schraw & Dennison, 1994; Zepeda & Nokes-Malach, 2021). The use of monitoring strategies is regarded as highly relevant within the learning process as they inform learners about difficulties in understanding and can help to identify measures to overcome these knowledge gaps (Vosniadou et al., 2021). Furthermore, empirical studies indicated that higher education students who regularly monitor their understanding show better performance (Broadbent & Poon, 2015; Credé & Phillips, 2011). This is also illustrated by training studies on monitoring strategy use that found performance effects on regular course exams in higher education (de Bruin, Kok, Lobbestael, & de Grip, 2017; Nietfeld, Cao, & Osborne, 2006).

Ideally, monitoring strategies are employed flexibly depending on the specific task requirements. Besides typical strategies that work for monitoring comprehension and retention of concepts, some tasks potentially require specific monitoring strategies. For example, to detect difficulties in understanding text, skimming a text could be useful while self-testing could be employed to monitor vocabulary knowledge. Therefore, empirical approaches should also consider situation-specific perspectives related to concrete learning tasks (e.g., Azevedo, 2009; Jansen, van Leeuwen, Janssen, Kester, & Kalz, 2016; McCardle & Hadwin, 2015; Rovers, Clarebout, Savelberg, de Bruin, & van Merriënboer, 2019).

When students monitor their understanding or performance, they usually differ in the quality of monitoring (Azevedo, 2009). Students who achieve a high quality of monitoring, use appropriate strategies and apply them in an elaborated and not only superficial way (Glogger, Schwonke, Holzäpfel, Nückles, & Renkl, 2012; Leutner, Leopold, & Den Elzen-Rump, 2007; Wirth & Leutner, 2008). For example, when engaging in self-testing, student A does this with a high degree of dedication. Student A writes answers down and compares personal answers to the content provided in the book. In contrast, student B looks up answers very quickly without deeply elaborating on the content. The quality of self-regulated learning strategies can be either rated by experts based on the sophistication of students' strategies (Glogger et al., 2012; Parkinson & Dinsmore, 2018; Wirth & Leutner, 2008) or estimated by the learners with regard to the target orientation, precision, and control of the application of a strategy, as established in the context of motivational regulation strategies by Engelschalk et al. (2017).

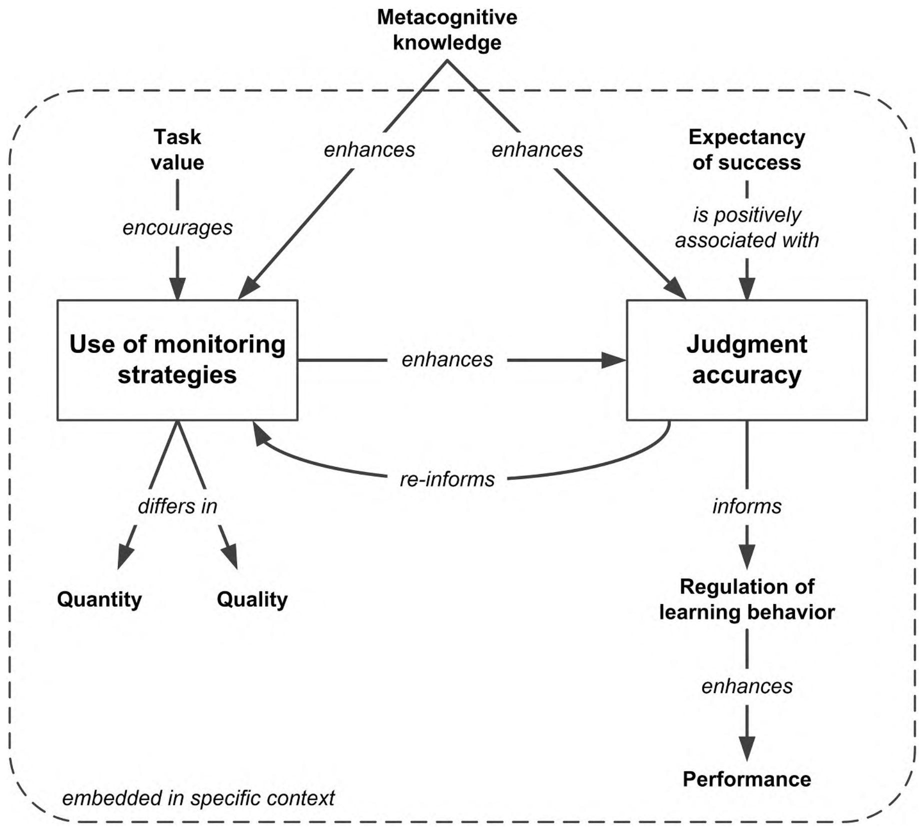

Fig. 1. Integrative theoretical model of metacognitive monitoring addressing the interplay of monitoring strategies and judgment accuracy as well as their common and specific determinants.

Consistent with the positive relationship between quality of cognitive strategy use and performance (Glogger et al., 2012; Haugwitz, Nesbit, & Sandmann, 2010; Leopold & Leutner, 2015), the quality of monitoring strategy use should be positively related with performance as well. Providing evidence for this, a qualitative study with higher education students (Ku & Ho, 2010) found that high performing students tended to use better quality monitoring strategies during critical thinking tasks.

1.2. Monitoring judgments and their quality

Overall, investigations in the strategy-oriented line of research have mostly been detached from the judgment-oriented line of research. Metacognitive frameworks (Flavell, 1979; Koriat, 2019) consider judgments as indicators for monitoring (Tarricone, 2011). Monitoring judgments are evaluations of personal performance or estimations about the correctness of task solutions. Several types of judgments should be distinguished (Hacker, Bol, & Keener, 2008; Nelson & Narens, 1990), especially depending on the point at which they are generated. Judgments before learning are called predictions, judgments during learning are so-called concurrent online judgments, and judgments after learning are called retrospective confidence judgments or postdictions. In addition, judgments differ in granularity. Global judgments refer to a whole test while item-specific judgments relate to a specific task. When learners generate judgments after learning or testing, they can base their judgments on experiences with the test material (Pierce & Smith, 2001). This is especially true for item-specific judgments that refer to single tasks and task-specific cues (Dunlosky & Lipko, 2007; Händel, de Bruin, & Dresel, 2020; Händel & Dresel, 2018). Consequently, item-specific judgments provide in-depth insights into already acquired knowledge or specific knowledge gaps (Händel, de Bruin, & Dresel, 2020; Händel & Fritzsche, 2016; Schraw, 2009a; Schraw, Kuch, & Gutierrez, 2013; Winne & Muis, 2011). This, in turn, provides more suitable information on which content needs to be studied further.

The quality of monitoring judgments is determined via judgment accuracy (Winne & Muis, 2011). For monitoring judgments, quality is indicated by several relative (e.g., gamma, sensitivity, specificity) and absolute (e.g., bias, absolute accuracy) accuracy scores (Bol & Hacker, 2012; Koriat, 2019). Research has convincingly demonstrated that higher education students differ in their accuracy of monitoring judgments with higher performing students showing higher accuracy. This applies not only for laboratory settings, as originally shown by Kruger and Dunning (1999), but also for field settings with graded exams (Bol & Hacker, 2001; Cao & Nietfeld, 2005; Händel, de Bruin, & Dresel, 2020). Judgment training in regular higher education courses has the potential to not only lead to more accurate judgments but also to performance improvements (Callender, Franco-Watkins, & Roberts, 2015; Händel, Harder, & Dresel, 2020). Hence, judgment accuracy seems to be clearly and positively related to performance. Learners who perform better seem to be better able to accurately judge their performance, which facilitates improvements in their knowledge.

1.3. Interrelations between monitoring strategies and monitoring judgments

Strategy-oriented research addresses monitoring strategies regarding all phases of self-regulated learning (Winne, 2010; Zimmerman, 2008). Judgment-oriented research, in contrast, focuses on evaluating personal knowledge or performance via monitoring judgments (Schraw, 2009b). Accordingly, these two lines of research address monitoring in different learning or performance phases (Dinsmore et al., 2008; Panadero, Jonsson, & Botella, 2017). Monitoring strategies are usually implemented during or after learning when learners reflect on their understanding or progress. As a result of a monitoring strategy (e.g., self-testing), students might generate a monitoring judgment.

So far, the interrelation between the use of monitoring strategies and the generation of monitoring judgments has not been empirically proven (Panadero et al., 2017). Only a few studies have investigated the use of monitoring strategies as well as monitoring judgments in the same participants regarding identical learning or testing material (Gidalevich & Kramarski, 2017; Schraw & Dennison, 1994; Sperling et al., 2004). This seems to be an important prerequisite in order to obtain more comprehensive insights into metacognitive monitoring. However, monitoring strategies are usually not well represented within most questionnaires and have only been assessed within a conglomerate of metacognitive strategies in general (Griffin et al., 2013). Training studies with approaches that integrate monitoring strategies and monitoring judgments have found differential performance effects. For example, in one study, training higher education students to engage in monitoring strategies led to better exam scores, but judgments were not related to exam scores (de Bruin et al., 2017). This provides a hint that monitoring strategies and monitoring judgments function differently. To understand the role of metacognitive monitoring during self-regulated learning and to consider metacognitive monitoring in all phases of self-regulation, an integrated approach of metacognitive monitoring is needed that considers monitoring strategies as well as monitoring judgments (Azevedo, 2009; Baars, Wijnia, de Bruin, & Paas, 2020; Dinsmore et al., 2008; Dunlosky & Thiede, 2013).

We propose an approach that integrates the quantity and quality of monitoring strategy use as well as judgment accuracy as separate factors (see Fig. 1). We assume that monitoring strategies and judgment accuracy are separate, yet highly related aspects of metacognitive monitoring. That is, a student who engages in monitoring strategies during learning, for example, via self-testing or explaining content to others, might thereby get insights into his or her understanding and accordingly make accurate judgments. This assumption is supported by the meta-analysis by Gutierrez de Blume (2021) indicating a moderate effect size of the influence of strategy training on judgment accuracy compared to a control group. Conversely, accurate judgments potentially lead to more or less intense use of monitoring strategies as indicated in a current study with undergraduates in a physics course (Morphew, 2021). Hence, we assume that metacognitive monitoring encompasses different but reciprocal processes.

1.4. *Common and specific determinants of monitoring strategies and judgment accuracy*

Metacognitive monitoring relates to several learner and task characteristics (Efklides, 2011). Motivation (Pintrich, 1999) and metacognitive knowledge (Dörrenbächer-Ulrich, Weißenfels, Russer, & Perels, 2021; Händel, Artelt, & Weinert, 2013) seem to be two particularly important factors for metacognition. In the framework of monitoring judgments, so-called "cues" are discussed as predictors (Koriat, 1997). These can relate to experiences during learning – so-called experience-based cues – or can be based on information gathered earlier – so-called information-based cues, according to Koriat, Nussinson, Bless, and Shaked (2008).

The importance of motivational components for self-regulated learning strategies has been indicated in several theoretical models on self-regulated learning (Boekaerts & Cascallar, 2006; Pintrich, 1999; Zimmerman, 2013). For example, Zimmerman (2008) emphasized the role of motivation in the forethought phase of self-regulated learning. The model refers to both expectancy-related factors (e.g., academic self-concept, expectancy of success) and value-related factors (e.g., intrinsic value, attainment value, goal orientation), which can be differentiated according to the expectancy-value theory of achievement motivation (Wigfield & Eccles, 2000). The importance of motivation for self-regulated learning strategies in general was also underpinned by a current review (Lim & Yeo, 2021). Consequently, and more specifically, motivational aspects should also be relevant for monitoring strategies.

Metacognitive knowledge (Flavell, 1979) seems to be essential for metacognitive monitoring. This relates to both the use of monitoring strategies and judgment accuracy (Negretti, 2021; Schraw, 1994, 1997). Procedural metacognitive strategy knowledge, that is, knowledge about effective and situation-adequate strategies may help to use appropriate, high quality monitoring strategies. For example, metacognitive strategy knowledge and the use of self-regulated learning strategies (i.e., control strategies) were positively interrelated across countries participating in a PISA study (Artelt & Schneider, 2015). In addition, students who have better knowledge on how to successfully monitor their understanding should be better able to judge their performance.

Besides these determinants of metacognitive monitoring, the proposed model in Fig. 1 displays assumed effects on regulation behavior and related student performance, which, however, are not investigated in the current studies.

1.5. *Aims of the studies*

The current work pursued two aims: First, to investigate the factorial structure of metacognitive monitoring and the interrelations of their factorial components; second, to model their potentially specific relationships with metacognitive knowledge and achievement motivation as important determinants of metacognitive monitoring. We report two studies that focused on metacognitive monitoring on two different levels of generalization. Study 1 addressed metacognitive monitoring on the global level of a whole subject and with regard to a graded final exam in this subject. Study 2 addressed situation-specific metacognitive monitoring regarding an ungraded midterm exam with specific content.

1.5.1. *Factorial structure of metacognitive monitoring*

We assumed that the three theoretically distinguished components of monitoring, namely quantity and quality of monitoring strategy use, and judgment accuracy can be separated empirically.[2] The hypothesis regarding the factorial structure of metacognitive monitoring is as follows:

H1. Quantity and quality of monitoring strategy use, and judgment accuracy are three individual factors of metacognitive monitoring that are positively interrelated.

1.5.2. *Common and specific determinants*

We assumed metacognitive knowledge to be a common determinant of the quantity and quality of monitoring strategy use as well as judgment accuracy.

H2. Students with higher metacognitive knowledge use more monitoring strategies, apply them more effectively, and generate more accurate monitoring judgments.

Furthermore, we expected value and expectancy of success to be specific determinants. Regarding the quantity and quality of monitoring strategy use, we suppose that they are particularly related to the value component of achievement motivation. Students who value their studies should use more monitoring strategies and apply them more effectively. In contrast, expectancy of success should be particularly related to judgment accuracy. In line with theoretical deliberations (Koriat, 1997) as well as current findings (Golke, Steininger, & Wittwer, 2022; Händel, de Bruin, & Dresel, 2020), students with high expectancies can base their judgments on earlier performance or effort as valid information-

[2] For monitoring judgments, we assume that the judgment itself provides information about the level of the judgment while accuracy scores provide information on their fit to actual performance, i.e., the quality of the judgments. However, accuracy scores are statistically dependent on the judgment level as judgment accuracy is calculated as the absolute difference of the performance level and the judgment level. Accordingly, it is not reasonable to statistically model this assumption.

based cues.

H3a. The value component of achievement motivation is associated with a more frequent (i.e., higher quantity) and a more elaborate (i.e., higher quality) use of monitoring strategies.

H3b. Expectancy of success is associated with higher judgment accuracy.

2. General method

2.1. Procedure

To provide evidence on the factorial structure and relationship of monitoring strategies and judgment accuracy, we conducted two studies focusing on metacognitive monitoring and its determinants. Both studies pursued an ecologically valid setting as all measurements referred to complex and relevant course content and exam preparation. Table 2 provides an overview of the implemented measurements.

In both studies, the assessment of metacognitive monitoring comprised questionnaire scales on the quantity and quality of monitoring strategy use, as well as monitoring judgments. Both studies focused on item-specific postdictions that display the most frequently used assessment procedure for monitoring judgments within different subjects, age groups, and settings (de Bruin et al., 2017; Händel, de Bruin, & Dresel, 2020; Kolić-Vehovec, Pahljina-Reinić, & Rončević Zubković, 2021; Rudolph, Niepel, Greiff, Goldhammer, & Kröner, 2017; Schnaubert & Bodemer, 2017). In addition, potential determinants of metacognitive monitoring were investigated. We implemented a metacognitive knowledge test with a focus on monitoring strategies as well as questionnaire scales regarding the value and expectancy of success as two components of achievement motivation.

Table 2
Overview of metacognitive and motivational measurements.

	Study 1: Global monitoring	Study 2: Situation-specific monitoring
Metacognitive monitoring		
Quantity of monitoring strategy use	Questionnaire scale by Wild and Schiefele (1994); Sample item: "In between, I would think about which parts I have already understood"; α = 0.64	Questionnaire scale by McCardle and Hadwin (2015); Sample item: "I asked myself if I was understanding the material"; α = 0.94
Quality of monitoring strategy use (Engelschalk et al., 2017)	Sample item: "I would make sure that I use the strategy correctly"; α = 0.84.	Sample item: "I made sure that I used the strategy correctly"; α = 0.93.
Monitoring judgments (Händel, Harder, & Dresel, 2020)	"Do you think your answer is correct?"; α = 0.71	"Do you think your answer is correct?"; α = 0.96
Potential correlates		
Expectancy of success (Rakoczy, Buff, & Lipowsky, 2005)	Sample item: "I think I will do well on the final exam"; α = 0.82	Sample item: "I think I will do well on the midterm exam"; α = 0.84
Value; attainment of achievement (Gaspard et al., 2015)	Sample item: "It is important to me to be good at the final exam"; α = 0.93	Sample item: "It is important to me to be good at the midterm exam"; α = 0.96
Metacognitive knowledge (Dörrenbächer-Ulrich et al., 2021)	Two scenarios of a metacognitive knowledge test, focusing on monitoring strategies. α = 0.86	Two scenarios of a metacognitive knowledge test, focusing on monitoring strategies. α = 0.91

Note. Items regarding the quantity and quality of monitoring strategy use and achievement motivation items had to be rated on a six-point Likert scale (1 = *not at all true* to 6 = *absolutely true*). Metacognitive judgments had to be rated via a dichotomous scale (0 = *no*, 1 = *yes*). Items of the metacognitive knowledge test had to be rated on a four-point Likert scale of usefulness (1 = *not at all useful* to 4 = *very useful*).

Both studies were conducted with undergraduate teacher education students from different courses at two separate German universities. Participants' privacy was protected; all data were anonymized, and participating students were not disadvantaged due to nonparticipation. Informed consent of the participants was obtained by virtue of survey completion.

Study 1 focused on the common scenario of exam preparation in higher education. Participating students were asked to think about a typical exam at the end of the term. They reported on how they would monitor their learning in such a scenario. Study 2 focused on situation-specific monitoring in an ungraded midterm exam. Students took part in the ungraded midterm exam and reported on how they monitored their learning.

2.2. Data analyses

To investigate the factorial structure and relationship of metacognitive monitoring components (H1), confirmatory factor analyses (CFA) were run. We performed three CFAs with items assessing monitoring strategies and judgment accuracy using the R package lavaan (Rosseel, 2012). First, a one-factor model with all items loading on one single factor was calculated. Next, a two-factor model separating monitoring strategies and judgment accuracy was calculated. Finally, the hypothesized three-factor model was calculated with items regarding quantity and quality of monitoring strategy use loading on two separate factors and judgment accuracy loading on a third factor. The χ^2-difference-test was used to compare the model fit of the three models. It is a robust test that compares the model fit of nested models, but with the limitation that it relies on the assumption that the larger model being compared is correctly specified (Pavlov, Shi, & Maydeu-Olivares, 2020; Yuan & Bentler, 2004).

To investigate hypotheses H2 and H3a-H3b, structural equation models (SEM) were calculated with the three monitoring factors simultaneously regressed on metacognitive knowledge, attainment value, and expectancy of success. First, a full model was specified with all paths estimated freely. In four further models, we set single regression paths as equal (i.e., paths of quantity of monitoring strategy use and judgment accuracy as well as paths of quality of monitoring strategy use and judgment accuracy – each for the value and the expectancy component of achievement motivation). We used χ^2-differences to test whether these constraints resulted in a significant drop in the model fit and thus indicate significantly different regression paths.

3. Study 1

3.1. Procedure

Study 1 focused on the common scenario of exam preparation in higher education. All study materials were paper-and-pencil, implemented within a regular psychology course. Students first completed a general knowledge test and respective item-specific judgments. This allowed us to assess students' judgment accuracy in the domain of psychology. Next, students reported on attainment value and expectancy of success regarding the psychology exam. Afterwards, students provided information on the quantity and quality of monitoring strategy use in prototypical exam preparation phases. Finally, students completed a metacognitive knowledge test.

3.2. Sample

Participants came from a mandatory psychology course for undergraduate teacher education students at a German university. All students enrolled in the course who were present at the first course session were invited to participate in the study. We did not exclude students from participation or pay them for participation. After providing study consent, 396 students voluntarily participated. The sample consisted of a high proportion of female (78.5 %) and freshmen (80.5 %) students. Students reported their GPA,[3] $M = 2.5$, $SD = 0.5$. Nonresponse missing rate ranged from 1.3 % to 7.6 % on the item level.

3.3. Measurements

The measurements were framed in the context of a common scenario of exam preparation in higher education. Table 2 provides an overview of the measurements.

3.3.1. Quantity of monitoring strategy use

To assess the frequency of monitoring strategy use, four questionnaire items of the LIST inventory (Wild & Schiefele, 1994), which is a German analogue of the MSLQ (Pintrich, Smith, Garcia, & McKeachie, 1991), were implemented.

3.3.2. Quality of monitoring strategy use

To assess the quality of monitoring strategy use, we applied an approach originally developed with regard to motivation regulation strategies (Engelschalk et al., 2017). Students described in their own words (open answer, not further analyzed) which monitoring strategy they would use when preparing for a final exam to monitor their understanding and progress. Next, students answered four items regarding the quality of their planned strategy. In accordance with Engelschalk et al. (2017), the items covered three quality aspects, namely target orientation, precision, and control. Quality was coded as the mean score of the closed-ended questionnaire items.[4]

3.3.3. Monitoring judgments

Item-specific judgments were collected in relation to a general knowledge test on psychology (ungraded test, 18 single choice items, one answer correct out of four options, $M = 64.6$, $SD = 12.6$). After completing each knowledge test item, students were asked to generate an item-specific monitoring judgment.

Absolute accuracy as a unipolar absolute measure[5] was chosen to operationalize judgment quality. It represents students' ability to accurately judge their item solution.

Scores close to one point for accurate monitoring and values close to zero indicate inaccurate judgments:

$$\text{Absolute accuracy} = 1 - \frac{1}{n}\sum_{i=1}^{n}|c_i - p_i|. \tag{1}$$

3.3.4. Attainment value

Attainment value was assessed as an indicator for the value component of achievement motivation because this facet is assumed to relate closely to monitoring learning progress and performance, while, for example, intrinsic value would be more relevant for motivational regulation but not for monitoring. We used the questionnaire scale attainment of achievement (Gaspard et al., 2015). The four items originally related to mathematics were adapted to performance in a psychology exam.

3.3.5. Expectancy of success

The expectancy component of achievement motivation was assessed via a four-item questionnaire scale developed by Rakoczy et al. (2005). Students were asked to rate their expectancy of success related to the final psychology exam at the very end of the semester.

3.3.6. Metacognitive knowledge

To assess metacognitive knowledge regarding monitoring strategies, a scenario-based knowledge test was implemented (Dörrenbächer-Ulrich et al., 2021).[6] Both scenarios describe the situation of a student preparing for an exam. Along with each scenario, six strategy options were provided. According to theoretical assumptions and expert evaluations, three of them are regarded as useful and three as less useful strategies. In contrast to the scales regarding quantity and quality of monitoring strategy use, the metacognitive knowledge test aims to assess students' knowledge of the usefulness of different strategies for specific learning situations rather than their personal behavior or preferences. The scoring of the test was applied via pair comparisons of each two provided strategies. In line with the procedure described by the authors, 18 pair comparisons of useful and less useful strategies were calculated. Thereby, a relative comparison of useful and less useful strategies was applied. Students reached one point for each pair comparison if they rated the more useful strategy more highly than the less useful strategy. The sum score was recoded so that "1" represents all pair comparisons correct and "0" represents no pair comparison correct.

3.4. Results

3.4.1. Factorial structure of metacognitive monitoring (H1)

The factorial structure and relationship of metacognitive monitoring components were analyzed via CFAs.[7] A comparison of fit indices of the three CFAs favors the theoretically assumed three-factor model with quantity of monitoring strategy use, quality of monitoring strategy use, and judgment accuracy as three separate latent factors (H1, cf. Table 3). The χ^2 difference test comparing the models were significant (Model 1 vs. Model 2: $\Delta\chi^2(1) = 61.8$, $p < .001$; Model 2 vs. Model 3: $\Delta\chi^2(1) = 60.9$, $p < .001$). As expected, monitoring could be distinguished in the quantity of monitoring strategy use ($M = 4.4$, $SD = 0.7$), the quality of monitoring strategy use ($M = 4.6$, $SD = 0.7$), and judgment accuracy ($M = 0.7$, $SD = 0.1$).

Fig. 2 displays the interrelations of the three latent factors. Quantity and quality of monitoring strategy use were positively correlated. Students who used more monitoring strategies used them more effectively. The correlation can be considered medium (Cohen, 1988). Neither quantity nor quality of monitoring strategy use was related with judgment accuracy.

3.4.2. Common and specific determinants (H2, H3a-H3b)

Metacognitive knowledge, attainment value, and expectancy of success were investigated as potential predictors for the quantity and quality of monitoring strategy use and judgment accuracy. Students' metacognitive knowledge was above the scale average ($M = 0.8$, $SD = 0.2$). Students valued the importance of good grades ($M = 5.0$, $SD = 0.9$) and showed moderate scores regarding expectancy of success ($M = 3.8$, $SD = 0.7$).

Structural equation modelling was performed to answer Hypotheses H2 and H3a-H3b. Fig. 3 shows the full structural equation model with all paths estimated freely ($\chi^2(15) = 112.1$).

Students' metacognitive knowledge predicted quantity as well as quality of monitoring strategy use but not judgment accuracy (i.e., H2 could be only partly confirmed). As expected, the attainment value

[3] GPA displays students' self-reported final high school grade ranging from 1 to 6 with lower values indicating better performance.

[4] The original instruction and items in the German language as well as the English translation can be seen in the electronic supplement.

[5] A unipolar measure with high values indicating high quality was needed to analyze judgment accuracy within the same analysis as quantity and quality of monitoring strategy use. Bias as a bipolar score, in contrast, would represent inaccurate judgments via both high and low values. In addition, to keep the scoring of judgment accuracy coherent with the assessment of monitoring strategy use, an absolute measure was chosen compared to relative scores like gamma.

[6] A sample item is available from the author of the test.

[7] Missing values were handled using the full-information maximum likelihood estimation (FIML) method.

Table 3
Model fit statistics for confirmatory factor analyses assessing quantity and quality of monitoring strategy use and judgment accuracy.

Model		χ^2	*df*	CFI	TLI	RMSEA	SRMR
Study 1: Global monitoring regarding exam preparation							
1	Monitoring	480.4	299	0.776	0.757	0.039	0.057
2	Monitoring strategies; judgment accuracy	418.6	298	0.851	0.838	0.032	0.054
3	Quantity of monitoring strategy use; quality of monitoring strategy use; judgment accuracy	357.7	296	0.924	0.917	0.023	0.050
Study 2: Monitoring regarding the ungraded midterm exam							
1	Monitoring	1487.5	350	0.517	0.479	0.120	0.137
2	Monitoring strategies; judgment accuracy	901.8	349	0.765	0.746	0.084	0.072
3	Quantity of monitoring strategy use; quality of monitoring strategy use; judgment accuracy	445.5	347	0.958	0.954	0.036	0.055

Note. χ^2 = Yuan–Bentler robust test statistic; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = Root-mean-square error of approximation; SRMR = Standardized rootmean-square residual.

predicted quantity and quality of monitoring strategy use. Students who valued achieving a good performance used monitoring strategies more frequently and more elaborately (H3a). In line with the assumptions, judgment accuracy was predicted only by the expectancy component but not by the value component of achievement motivation (H3b). Students with a higher expectancy of success judged their performance in the general knowledge test more accurately than students with a lower expectancy of success. In addition, expectancy of success predicted the quantity of monitoring strategy use.

Table 4 shows the results of the model comparisons via the χ^2 differences test. A significant difference would indicate that the model with model constraints fits the data significantly worse than the full model, which means that they are specific determinants.

Significant differences could be found regarding attainment value, which was less related to judgment accuracy than to quantity and quality of monitoring strategy use, confirming H3a. Students who valued good grades engaged in more and more elaborated monitoring strategies but did not show higher judgment accuracy. In contrast to H3b, no significant differences for expectancy of success were found in the direction or strength of effects on the separate aspects of metacognitive monitoring.

3.5. *Discussion of Study 1*

Study 1 used the scenario of exam preparation in higher education to investigate the factorial structure of metacognitive monitoring. It provided evidence that metacognitive monitoring is composed of quantity of monitoring strategy use, quality of monitoring strategy use, and judgment accuracy (H1). Overall, the three factors were rather weakly correlated with each other—the lack of correlations with judgment accuracy was especially surprising against the background of theoretically highly intertwined facets of metacognitive monitoring. This is particularly striking against the background that both monitoring strategies and monitoring judgments related to the same object of monitoring, namely monitoring performance. The assumed common and specific determinants of different aspects of metacognitive monitoring could only be partly confirmed (H2, H3a-H3b).

Despite keeping the test format and test content coherent in Study 1, quantity and quality of monitoring strategy use related to a different performance situation (a graded final exam) than the monitoring judgments (an ungraded performance test). This might be a possible explanation for the small and non-significant correlations between monitoring strategy use and judgment accuracy in Study 1. In addition, students might differ in their monitoring behavior between different exams (e.g., with regard to subject, time available, or expected exam difficulty). Consequently, we implemented a specific scenario in Study 2. We contextualized all measures within one specific situation, namely an ungraded midterm exam.[8]

4. Study 2

4.1. *Procedure*

Study 2 used a situation-specific approach: Metacognitive monitoring was assessed regarding an ungraded midterm exam, which was announced ten days prior to administration. Students were informed that the test presented a good learning opportunity to review previous lecture content and to become familiar with the test format of the course exam. Students were not explicitly instructed on how to prepare for the test, but it was recommended that they study the same way as they would for a course exam. With this procedure, we aimed to achieve a situation-specific assessment of students' monitoring behavior and potential determinants.

Study 2 was conducted online via the survey tool Unipark. The instruments were administered as follows: First, students reported on attainment value and expectancy of success. Second, students reported the degree and quality of their engagement in monitoring strategies during the self-paced preparation phase. Afterwards, students completed the ungraded midterm exam and generated item-specific judgments. Finally, students completed the metacognitive knowledge test.

4.2. *Sample*

Participants came from a mandatory psychology course for undergraduate teacher education students at a different German university. Students were invited to participate in an ungraded midterm exam and the additional survey questions. Due to course relatedness, we did not exclude students from participation. As an incentive for complete participation in the survey, students could participate in a raffle offering one of three €50 gift cards. After providing study consent, 225 students (76.9 % female, 71.6 % enrolled in the second study year, GPA, $M = 2.3$, $SD = 0.6$) voluntarily participated. Full data were available as students had to respond to all items in the online survey tool.

4.3. *Measurements*

Study 2 adapted the measuring instruments from Study 1 to the situation-specific context. First, while all questionnaire items in Study 1 were framed in the context of general exam preparation ("Imagine you are preparing for a psychology exam"), Study 2 focused on a concurrent situation of an ungraded midterm exam ("Think about your personal preparation for this midterm exam"). Accordingly, all items regarding the preparation phase were formulated in the past tense. Except for these adaptions, quality of monitoring strategies, metacognitive knowledge, attainment value, and expectancy of success were assessed via the same

[8] Students prepared for the ungraded midterm exam, they reflected on their personal value and expectancy of success as well as on how they monitored their learning. Students filled in the ungraded midterm exam and judged their performance for each item.

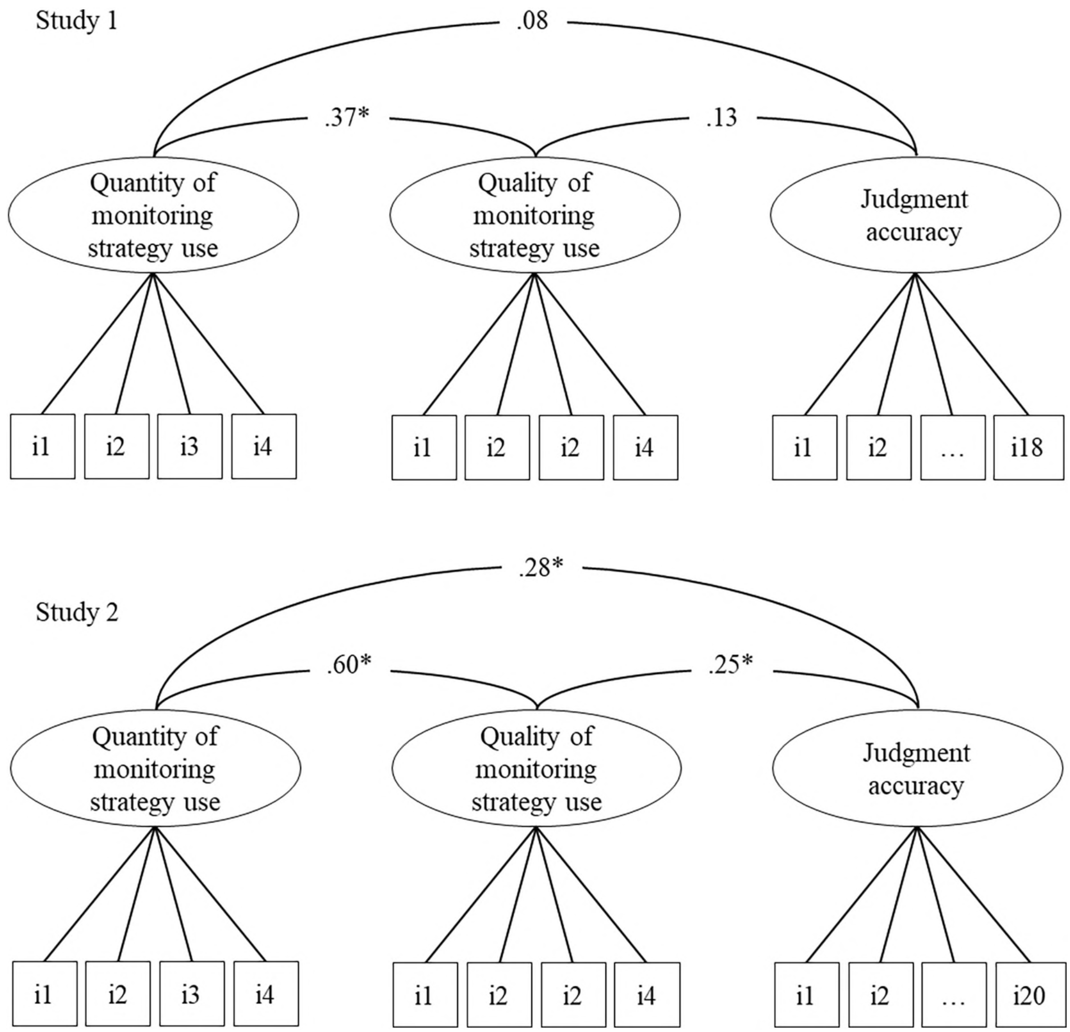

Fig. 2. Interrelations of the three latent factors (* $p < .05$) in Study 1 (global metacognitive monitoring) and Study 2 (situation-specific metacognitive monitoring).

scales as in Study 1. The scale quantity of monitoring strategy use was replaced by a four-item questionnaire scale from McCardle and Hadwin (2015) due to rather low internal consistency of the questionnaire scale in Study 1. Finally, monitoring judgments were collected in relation to the ungraded midterm exam (20 multiple choice items, one to four answers correct out of four options, $M = 62.6$ % correct, $SD = 8.1$). Table 2 provides an overview of metacognitive and motivational measurements, including sample items.

4.4. Results

4.4.1. Factorial structure of metacognitive monitoring (H1)

As in Study 1, three CFAs were calculated to test Hypothesis H1 (cf. Table 3, lower part). The comparison of fit indices again favors the theoretically assumed three-factor model with quantity of monitoring strategy use, quality of monitoring strategy use, and judgment accuracy as three separate latent factors. The χ^2 difference tests comparing the models are each significant (Model 1 vs. Model 2: $\Delta\chi^2(1) = 585.7$, $p < .001$; Model 2 vs. Model 3: $\Delta\chi^2(1) = 456.3$, $p < .001$). Additionally, for the situation-specific assessment, students' metacognitive monitoring could be distinguished in the quantity of monitoring strategy use ($M = 4.0$, $SD = 1.2$), the quality of monitoring strategy use ($M = 3.8$, $SD = 1.1$), and judgment accuracy ($M = 0.7$, $SD = 0.1$), thereby confirming Hypothesis H1.

The latent correlation of quantity and quality of monitoring strategy use (see Fig. 2, bottom part) can be considered strong (Cohen, 1988). Still, quantity and quality of monitoring strategy use could be separated empirically and displayed two separate aspects of monitoring—as the statistical comparison of the respective models indicated. In addition, quantity and quality of monitoring strategy use can each be separated from judgment accuracy, to which both factors were weakly correlated.

4.4.2. Common and specific determinants (H2, H3a-H3b)

The relation of metacognitive monitoring with metacognitive knowledge and achievement motivation was investigated. Again, students' metacognitive knowledge was descriptively above average ($M = 0.8$, $SD = 0.3$). Achievement motivation, however, seemed to be lower than in Study 1 (attainment value: $M = 3.2$, $SD = 1.2$; expectancy of success: $M = 3.1$, $SD = 0.7$).

Fig. 3 (bottom part) displays the full model ($\chi^2(15) = 227.9$) and shows that all paths were significant—except for judgment accuracy not being regressed on the value component.

Table 4 shows the comparison of model fits when specific pairs of regression paths were held equal. This analysis indicated significantly different paths with monitoring strategies and judgment accuracy for attainment value. Attainment value only predicted the quantity and quality of monitoring strategy use but not judgment accuracy (H3a). Beyond that, no further specific predictors were identified (i.e., H3b could not be confirmed). Expectancy of success was equally important for the quantity and quality of monitoring strategy use as well as for judgment accuracy.

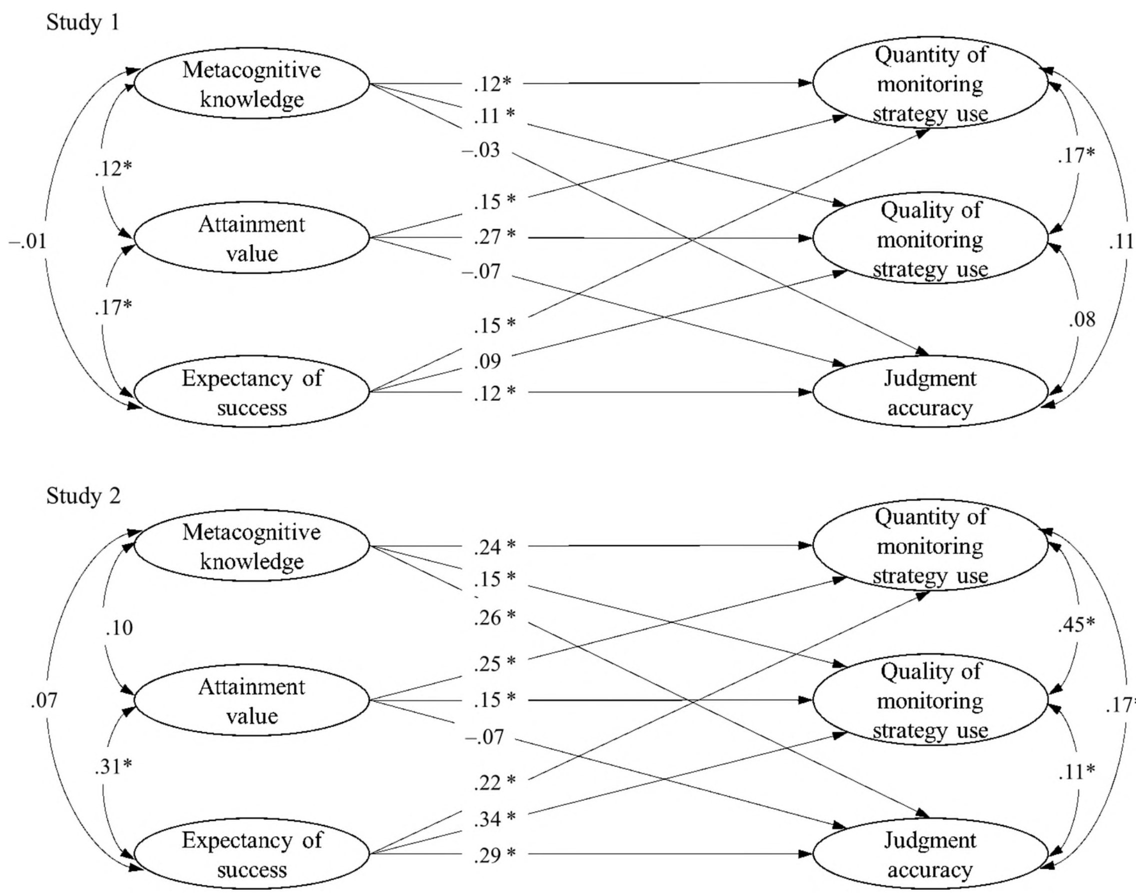

Fig. 3. Structural equation models for Study 1 (global metacognitive monitoring) and Study 2 (situation-specific metacognitive monitoring), presented are standardized regression coefficients (* $p < .05$).

Table 4
Chi-square difference results comparing the full model with freely estimated regression paths to models with each two regression paths set equal.

Predictor	Regression paths set equal for quantity of monitoring strategy use and accuracy	Regression paths set equal for quality of monitoring strategy use and accuracy
	$\Delta\chi$ ($\Delta df = 1$)	$\Delta\chi$ ($\Delta df = 1$)
Study 1: Global monitoring regarding exam preparation		
Attainment value	7.8**	19.5***
Expectancy of success	0.2	0.1
Study 2: Monitoring regarding the ungraded midterm exam		
Attainment value	15.1***	7.1**
Expectancy of success	0.7	0.3

Note. *** $p < .001$, ** $p < .01$, * $p < .05$.

4.5. *Discussion of Study 2*

Study 2 was embedded in the concrete situation of preparing for and taking part in an ungraded midterm exam in higher education. It confirmed the factorial structure of metacognitive monitoring, composed of the three factors: quantity of monitoring strategy use, quality of monitoring strategy use, and judgment accuracy. It also revealed common and specific determinants for the different aspects of metacognitive monitoring.

5. General discussion

The present studies pursued an integrated approach to studying metacognitive monitoring. So far, different aspects of metacognitive monitoring, namely monitoring strategies (Winne & Hadwin, 2008) and judgments (Koriat, 2019), have only been studied separately from each other (Griffin et al., 2013). This is to the best of our knowledge the first study with a comprehensive approach simultaneously modelling the quantity and quality of monitoring strategy use as well as judgment accuracy. This integrated perspective is reflected in the proposed theoretical model (see Fig. 1). The few earlier studies that implemented metacognitive strategies as well as monitoring judgments did not explicitly measure monitoring strategies but typically analyzed them as part of larger measurements (Griffin et al., 2013). Moreover, previous questionnaire studies were restricted to the quantity aspect of monitoring strategy use. The factorial structure of metacognitive monitoring was tested across two different levels of generalization—global monitoring behavior regarding graded final exams and situation-specific monitoring regarding the preparation of a concrete ungraded midterm exam. In addition, the current two studies modelled the influence of metacognitive knowledge, attainment value, and expectancy of success simultaneously on different metacognitive monitoring factors.

5.1. *Factorial structure of metacognitive monitoring (H1)*

Across both studies, the results confirmed the proposed theoretical model for metacognitive monitoring provided in Fig. 1, composed of quantity and quality of monitoring strategy use as well as judgment accuracy. The assumption could be confirmed that monitoring strategies and judgment accuracy individually contribute to metacognitive monitoring (H1). The lower fit statistics for the alternative one- and two-factorial models clearly indicate that judgment accuracy is a separate factor of metacognitive monitoring in addition to the quantity and quality of monitoring strategy use. Confirming the multifactorial structure of metacognitive monitoring across both studies represents a key result of this research.

Small correlations in Study 1 and moderate correlations in Study 2 indicated that quantity and quality of monitoring strategy use were related to each other. In contrast, a small relationship of quantity and quality of monitoring strategy use with monitoring accuracy was found only in Study 2 but not in Study 1. This finding extends the scarce body of research regarding relationships of different measures of metacognitive strategies (Sperling et al., 2004), which was limited to a general assessment, small samples and heterogeneous results. Especially in Study 2 with the situation-specific assessment, students who implemented monitoring strategies more frequently and effectively showed higher judgment accuracy. Students who do not sufficiently engage in monitoring strategies during the learning process potentially have difficulties in recognizing their knowledge gaps, and, consequently, continue with rather ineffective learning behavior (Nugteren, Jarodzka, Kester, & Van Merriënboer, 2018).

Given the close theoretical overlap of monitoring strategies and monitoring judgments, quite large correlations between the factors were expected. A possible explanation for the – at most – low correlations between monitoring strategies and judgment accuracy in our studies is the situation-specificity of metacognitive monitoring (Azevedo, 2009; Roth, Ogrin, & Schmitz, 2015; Rovers et al., 2019). This can also have implications for the relationship between metacognitive monitoring and performance. For example, in the meta-analysis on self-regulated learning in primary and secondary education by Dent and Koenka (2015), only comprehension monitoring strategies but not general monitoring strategies were found to be associated with performance. Still, even in Study 2 focusing on a situation-specific approach, correlations between monitoring strategies and judgment accuracy were only small. Overall, the surprisingly low correlations clearly support the assumption of separate aspects of monitoring and call for future research.

In addition to the main finding that monitoring strategies and judgment accuracy are distinct components, the current two studies provided information on the distinction between quantity and quality of monitoring strategy use. The current research successfully adapted an efficient approach to assess strategy quality from the field of motivational regulation in higher education students (Engelschalk et al., 2017). This allowed us to distinguish between quantity and quality of monitoring strategy use (Wirth & Leutner, 2008). Regarding the global monitoring behavior during exam preparation in Study 1, quantity and quality of monitoring strategy use were correlated to a small degree. This is comparable to findings in the field of motivational regulation in higher education students (Steuer, Engelschalk, Eckerlein, & Dresel, 2019). Also in line with research on motivational regulation (Eckerlein et al., 2019) is the stronger relationship of quantity and quality of monitoring strategy use when referring to a concrete learning situation that students had just passed—preparation for an ungraded midterm exam in Study 2. Students differed in their application of monitoring strategies—not only regarding whether or how often they monitored their learning progress but also how precisely they monitored it. Hence, research on metacognitive monitoring may clearly distinguish aspects of quantity and quality of monitoring strategy use in the future.

5.2. *Common and specific determinants of metacognitive monitoring (H2 and H3a-H3b)*

In both studies, students who scored higher on the metacognitive knowledge test consistently reported using more monitoring strategies as well as higher quality monitoring strategies (H2). This is in line with our assumptions based on earlier research with PISA data (Artelt & Schneider, 2015). Interestingly, metacognitive strategy knowledge and judgment accuracy were related to each other only in Study 2 with the situation-specific assessment. To explain the heterogeneous results, future research is needed to clarify the relevance of metacognitive knowledge for judgment accuracy. It is likely that students do not only use strategy knowledge when making monitoring judgments but also refer to person- or task-related metacognitive knowledge (i.e., the dedication with which they prepare for an exam, how successful they were in the past, how difficult they think the test is; Flavell, 1979). This, however, was not assessed in the current studies.

Both quantity and quality of monitoring strategy use was predicted by value and expectancy components of achievement motivation, which is in line with previous research on quantity aspects of monitoring (Dresel & Haugwitz, 2006). Across the two studies, students who valued personal success in the midterm/final exam implemented monitoring strategies more often and more effectively (H3a). Across both studies, the influence of attainment value differed for monitoring strategies and judgment accuracy. This provides important insights into factors relevant for different aspects of metacognitive monitoring. Students who valued good performance used more and better quality monitoring strategies. In contrast, high attainment value did not necessarily lead to accurate judgments. Conversely, on top of the assumed influence of expectancy of success on judgment accuracy (H3b), we found additional effects of expectancy of success on quantity and quality of monitoring strategy use.

Overall, the results are predominantly in line with our proposed model and indicate that monitoring strategies and judgment accuracy seem to have at least partly different motivational determinants. The results confirm the role of achievement motivation in the forethought phase of self-regulated learning, as proposed by Zimmerman (2008). Furthermore, the results extend the literature by distinguishing the differential relevance of value and expectancies for different components of metacognitive monitoring. Still, due to the cross-sectional design of both studies, results need to be interpreted with caution.

5.3. *Limitations and directions for future research*

Results are limited to samples of voluntarily participating teacher education students in the subject of psychology. Based on findings on the subject-specificity of metacognitive monitoring (Dentakos, Saoud, Ackerman, & Toplak, 2019; Greene et al., 2015; Vanderstoep, Pintrich, & Fagerlin, 1996), replication studies with other study subjects and additional types of tests or answer formats are needed. Moreover, it cannot be ruled out that specific groups of students—for example, high performing or highly motivated students—took part in the studies with greater probability. However, while self-selection may have influenced the mean levels of the individual constructs, it should not necessarily influence the factorial structure of metacognitive monitoring, which is the focus of this research. In addition, students supposedly alter their monitoring behavior with study experience (Pokay & Blumenfeld, 1990). Hence, research can benefit from the consideration of students' performance levels as well as their study experiences. Moreover, research is needed that replicates and generalizes the current findings to other types of monitoring strategies (e.g., monitoring of goal achievement, monitoring of cognitive strategy use; cf., Azevedo, Behnagh, Duffy, Harley, & Trevors, 2012) and monitoring judgments (e.g., predictions, postdictions but also judgments of learning vs. judgments of understanding; cf., Pilegard & Mayer, 2015).

As the studies focused on either global monitoring regarding graded final exams or an ungraded specific performance situation, further research should investigate whether results transfer to situation-specific high stakes scenarios where students potentially learn and thereby monitor more extensively. Students' monitoring in low- versus high-stakes situations could be influenced by utility and cost in different ways (Karabenick, Berger, Ruzek, & Schenke, 2021). Monitoring in low-stakes situations might be relevant to review one's current state of knowledge while monitoring regarding high-stakes situations might be highly relevant regarding study success.

Based on the findings on the factorial structure of metacognitive monitoring, the next step would be to investigate how monitoring strategies and monitoring judgments are reciprocally intertwined. This can be studied via modelling the causal relationship of monitoring strategies and monitoring judgments. As indicated in Fig. 1, we assume a reciprocal relationship between monitoring strategies and judgment accuracy. While it is theoretically sound to assume that high quantity and quality of monitoring strategy use lead to higher judgment accuracy, it can be assumed that the reverse path from judgment accuracy to monitoring strategies depends on the level of performance. While accurate judgments of high performance might serve as an indicator for a successful learning process and consequently reduces the quantity of monitoring strategy use, accurate judgments of low performance might indicate that quantity and quality of monitoring strategy use should be intensified. To investigate this potential relationship, process-oriented research designs with situation-specific assessment are needed. Moreover, we suggest that future research emphasizes not only the situation that requires monitoring (e.g., low- vs. high-stakes) but also the phase in the cycle of self-regulated learning (Zimmerman, 1998). Monitoring strategies might not only be applied by students in the preparatory phase for an exam but also during the exam itself. For example, students might check whether personal answers can be correct via processes of elimination. Similarly, monitoring judgments might not only be generated during test taking but also before (Ariel, Karpicke, Witherby, & Tauber, 2021; Jordano & Touron, 2018).

Finally, to investigate the significance of metacognitive monitoring during self-regulated learning, it seems fruitful to evaluate the common and specific effects of training approaches to enhance the use of monitoring strategies and judgment accuracy. Earlier research in school and higher education settings indicated positive effects of isolated training approaches for self-regulated learning strategies (Donker, de Boer, Kostons, Dignath van Ewijk, & van der Werf, 2014; Dresel & Haugwitz, 2008) and for monitoring judgments (Händel, Harder, & Dresel, 2020; Miller & Geraci, 2011; Naujoks, Harder, & Händel, 2022) as well as positive effects of combined training approaches (de Bruin et al., 2017; Huff & Nietfeld, 2009).

5.4. Conclusion

The current work proposed an integrated theoretical understanding and model of metacognitive monitoring as well as its determinants and provided evidence for the proposed integrated approach. The studies indicated that the quantity and the quality of monitoring strategy use, as well as judgment accuracy display three distinct components of monitoring that interrelate to a remarkably low degree. The three factors of metacognitive monitoring have common determinants, namely metacognitive knowledge and expectancy of success, as well as specific determinants, namely attainment value. Future research considering metacognitive monitoring from an integrated perspective will add to our understanding of how different aspects of metacognitive monitoring are related and how they contribute to regulation and performance.

Acknowledgements

We thank Laura Dörrenbächer-Ulrich for providing us with the test and scoring procedure of the metacognitive knowledge test.

We thank Nick Naujoks and Susanne Schober for their support in the administration of data collection of Study 1.

References

Ariel, R., Karpicke, J. D., Witherby, A. E., & Tauber, S. K. (2021). Do judgments of learning directly enhance learning of educational materials? *Educational Psychology Review, 33*(2), 693–712. https://doi.org/10.1007/s10648-020-09556-8

Artelt, C., & Schneider, W. (2015). Cross-country generalizability of the role of metacognitive knowledge in students' strategy use and reading competence. *Teachers College Record, 117*, 1–32.

Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition Learning, 4*, 87–95. https://doi.org/10.1007/s11409-009-9035-7

Azevedo, R., Behnagh, R., Duffy, M., Harley, J., & Trevors, G. (2012). Metacognition and self-regulated learning in student-centered leaning environments. In D. Jonassen, & S. Land (Eds.), *Theoretical foundations of student-centered learning environments* (Vol. 2, pp. 171–197). Routledge.

Baars, M., Wijnia, L., de Bruin, A., & Paas, F. (2020). The relation between students' effort and monitoring judgments during learning: A meta-analysis. *Educational Psychology Review, 32*(4), 979–1002. https://doi.org/10.1007/s10648-020-09569-3

Boekaerts, M. (1999). Motivated learning: Studying student * situation transactional units. *European Journal of Psychology of Education, 14*(1), 41–55. https://doi.org/10.1007/bf03173110

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review, 18*(3), 199–210. https://doi.org/10.1007/s10648-006-9013-4

Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *The Journal of Experimental Education, 69*, 133–151. https://doi.org/10.1080/00220970109600653

Bol, L., & Hacker, D. J. (2012). Calibration research: Where do we go from here? *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00229

Bong, M. (1997). *Congruence of measurement specificity on relations between academic self-efficacy, effort, and achievement indexes*. Chicago: AERA.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Callender, A. A., Franco-Watkins, A. M., & Roberts, A. S. (2015). Improving metacognition in the classroom through instruction, training, and feedback. *Metacognition and Learning, 11*, 215–235. https://doi.org/10.1007/s11409-015-9142-6

Cao, L., & Nietfeld, J. L. (2005). Judgment of learning, monitoring accuracy, and student performance in the classroom context. *Current Issues in Education, 8*(4), 1–24.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2. ed.). Academic Press.

Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: Systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning, 15*(2), 155–213. https://doi.org/10.1007/s11409-020-09222-y

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning and Individual Differences, 21*, 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

de Bruin, A. B. H., Kok, E. M., Lobbestael, J., & de Grip, A. (2017). The impact of an online tool for monitoring and regulating learning at university: Overconfidence, learning strategy, and personality. *Metacognition and Learning, 12*, 21–43. https://doi.org/10.1007/s11409-016-9159-5

Dent, A. L., & Koenka, A. C. (2015). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*, 425–474. https://doi.org/10.1007/s10648-015-9320-8

Dentakos, S., Saoud, W., Ackerman, R., & Toplak, M. E. (2019). Does domain matter? Monitoring accuracy across domains. *Metacognition and Learning, 14*(3), 413–436. https://doi.org/10.1007/s11409-019-09198-4

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review, 20*, 391–409. https://doi.org/10.1007/s10648-008-9083-6

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review, 11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002

Dörrenbächer-Ulrich, L., Weißenfels, M., Russer, L., & Perels, F. (2021). Multimethod assessment of self-regulated learning in college students: Different methods for different components? *Instructional Science, 49*, 137–163. https://doi.org/10.1007/s11251-020-09533-2

Dresel, M., & Haugwitz, M. (2006). The relationship between cognitive abilities and self-regulated learning: Evidence for interactions with academic self-concept and gender. *High Ability Studies, 16*, 201–218. https://doi.org/10.1080/13598130600618066

Dresel, M., & Haugwitz, M. (2008). A computer based training approach to foster motivation and self-regulated learning. *Journal of Experimental Education, 77*, 3–20. https://doi.org/10.3200/JEXE.77.1.3-20

Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., Jöstl, G., Klug, J., Roth, A., Wimmer, B., & Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: Structural model and indications drawn from expert interviews. *Studies in Higher Education, 40*, 454–470. https://doi.org/10.1080/03075079.2015.1004236

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension. *Current Directions in Psychological Science, 16*, 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: Why understanding students' judgments can improve their achievement. *Learning and Instruction, 24*, 58–61. https://doi.org/10.1016/j.learninstruc.2012.05.002

Eckerlein, N., Roth, A., Engelschalk, T., Steuer, G., Schmitz, B., & Dresel, M. (2019). The role of motivational regulation in exam preparation: Results from a standardized diary study. *Frontiers in Psychology, 10*, 81. https://doi.org/10.3389/fpsyg.2019.00081

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*, 6–25. https://doi.org/10.1080/00461520.2011.538645

Engelschalk, T., Steuer, G., & Dresel, M. (2017). Quantity and quality of motivational regulation among university students. *Educational Psychology, 37*, 1154–1170. https://doi.org/10.1080/01443410.2017.1322177

Flavell, J. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 35–60). Academic Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34*, 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*, 663–677. https://doi.org/10.1037/edu0000003

Gidalevich, S., & Kramarski, B. (2017). Guidance for metacognitive judgments: A thinking-aloud analysis in math problem solving. *Hellenic Journal of Psychology, 14* (2), 83–113.

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology, 104*, 452–468. https://doi.org/10.1037/a0026683

Golke, S., Steininger, T., & Wittwer, J. (2022). What makes learners overestimate their text comprehension? The impact of learner characteristics on judgment bias. *Educational Psychology Review*. https://doi.org/10.1007/s10648-022-09687-0

Greene, J. A., & Azevedo, R. (2007). A theoretical review of winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*(3), 334–372. https://doi.org/10.3102/003465430303953

Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C., & McVea, M. (2015). Domain-specificity of self-regulated learning processing in science and history. *Contemporary Educational Psychology, 42*, 111–128. https://doi.org/10.1016/j.cedpsych.2015.06.001

Griffin, T. D., Wiley, J., & Salas, C. R. (2013). Supporting effective self-regulated learning: The critical role of monitoring. In R. Azevedo, & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 19–34). Springer.

Gutierrez de Blume, A. P. (2021). Calibrating calibration: A meta-analysis of learning strategy instruction interventions to improve metacognitive monitoring accuracy. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000674

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429–455). Routledge.

Händel, M., Artelt, C., & Weinert, S. (2013). Assessing metacognitive knowledge: Development and evaluation of a test instrument. *Journal of Educational Research Online, 5*, 162–188.

Händel, M., de Bruin, A. B. H., & Dresel, M. (2020). Individual differences in local and global metacognitive judgments. *Metacognition and Learning, 15*, 51–75. https://doi.org/10.1007/s11409-020-09220-0

Händel, M., & Dresel, M. (2018). Confidence in performance judgment accuracy: The unskilled and unaware effect revisited. *Metacognition and Learning, 13*, 265–285. https://doi.org/10.1007/s11409-018-9185-6

Händel, M., & Fritzsche, E. S. (2016). Unskilled but subjectively aware: Metacognitive monitoring ability and respective awareness in low-performing students. *Memory & Cognition, 44*, 229–241. https://doi.org/10.3758/s13421-015-0552-0

Händel, M., Harder, B., & Dresel, M. (2020). Enhanced monitoring accuracy and test performance: Incremental effects of judgment training over and above repeated testing. *Learning and Instruction, 65*. https://doi.org/10.1016/j.learninstruc.2019.101245

Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences, 20*, 536–543. https://doi.org/10.1016/j.lindif.2010.04.004

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*, 161–176. https://doi.org/10.1007/s11409-009-9042-8

Jansen, R. S., van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2016). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education, 29*, 6–27. https://doi.org/10.1007/s12528-016-9125-x

Jordano, M. L., & Touron, D. R. (2018). How often are thoughts metacognitive? Findings from research on self-regulated learning, think-aloud protocols, and mind-wandering. *Psychonomic Bulletin & Review, 25*(4), 1269–1286. https://doi.org/10.3758/s13423-018-1490-1

Kaplan, A. (2008). Clarifying metacognition, self-regulation, and self-regulated learning: What's the purpose? *Educational Psychology Review, 20*, 477–484. https://doi.org/10.1007/s10648-008-9087-2

Karabenick, S. A., Berger, J.-L., Ruzek, E., & Schenke, K. (2021). Strategy motivation and strategy use: Role of student appraisals of utility and cost. *Metacognition and Learning, 16*(2), 345–366. https://doi.org/10.1007/s11409-020-09256-2

Kolić-Vehovec, S., Pahljina-Reinić, R., & Rončević Zubković, B. (2021). Effects of collaboration and informing students about overconfidence on metacognitive judgment in conceptual learning. *Metacognition and Learning*. https://doi.org/10.1007/s11409-021-09275-7

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A. (2019). Confidence judgments: The monitoring of object-level and same-level performance. *Metacognition and Learning, 14*, 463–478. https://doi.org/10.1007/s11409-019-09195-7

Koriat, A., Nussinson, R., Bless, H., & Shaked, N. (2008). Information-based and experience-based metacognitive judgments: Evidence from subjective confidence. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of memory and metamemory* (pp. 117–135). Psychology Press.

Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22*, 121–132. https://doi.org/10.1016/j.learninstruc.2011.08.004

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Ku, K. Y. L., & Ho, I. T. (2010). Metacognitive strategies that enhance critical thinking. *Metacognition and Learning, 5*(3), 251–267. https://doi.org/10.1007/s11409-010-9060-6

Lajoie, S. P. (2008). Metacognition, self regulation, and self-regulated learning: A rose by any other name? *Educational Psychology Review, 20*(4), 469–475. https://doi.org/10.1007/s10648-008-9088-1

Lan, W. (2005). Self-monitoring and its relationship with educational level and task importance. *Educational Psychology, 25*(1), 109–127. https://doi.org/10.1080/0144341042000294921

Leopold, C., & Leutner, D. (2015). Improving students' science text comprehension through metacognitive self-regulation when applying learning strategies. *Metacognition and Learning, 10*, 313–346. https://doi.org/10.1007/s11409-014-9130-2

Leutner, D., Leopold, C., & Den Elzen-Rump, V. (2007). Self-regulated learning with a text-highlighting strategy: A training experiment. *Zeitschrift für Psychologie, 215*, 174–182. https://doi.org/10.1027/0044-3409.215.3.174

Lim, S. L., & Yeo, K. J. (2021). The relationship between motivational constructs and self-regulated learning: A review of literature. *International Journal of Evaluation and Research in Education, 10*, 330–335. https://doi.org/10.11591/ijere.v10i1.21006

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning, 10*, 43–75. https://doi.org/10.1007/s11409-014-9132-0

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review, 15*, 174–179. https://doi.org/10.3758/PBR.15.1.174

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition and Learning, 6*, 303–314. https://doi.org/10.1007/s11409-011-9083-7

Morphew, J. W. (2021). Changes in metacognitive monitoring accuracy in an introductory physics course. *Metacognition and Learning, 16*(1), 89–111. https://doi.org/10.1007/s11409-020-09239-3

Naujoks, N., Harder, B., & Händel, M. (2022). Testing pays off twice: Potentials of practice tests and feedback regarding exam performance and judgment accuracy. *Metacognition and Learning*. https://doi.org/10.1007/s11409-022-09295-x

Negretti, R. (2021). Metacognition in academic writing: Learning dimensions. In H. Mohebbi, & C. Coombe (Eds.), *Research questions in language education and applied linguistics: A reference guide* (pp. 243–247). Springer.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26*, 125–141. https://doi.org/10.1016/S0079-7421(08)60053-5

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*, 159–179. https://doi.org/10.1007/s10409-006-9595-6

Nugteren, M. L., Jarodzka, H., Kester, L., & Van Merriënboer, J. J. G. (2018). Self-regulation of secondary school students: Self-assessments are inaccurate and insufficiently used for learning-task selection. *Instructional Science, 46*(3), 357–381. https://doi.org/10.1007/s11251-018-9448-2

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Parkinson, M. M., & Dinsmore, D. L. (2018). Multiple aspects of high school students' strategic processing on reading outcomes: The role of quantity, quality, and conjunctive strategy use. *British Journal of Educational of Psychology, 88*(1), 42–62. https://doi.org/10.1111/bjep.12176

Pavlov, G., Shi, D., & Maydeu-Olivares, A. (2020). Chi-square difference tests for comparing nested models: An evaluation with non-normal data. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(6), 908–917. https://doi.org/10.1080/10705511.2020.1717957

Pierce, B. H., & Smith, S. M. (2001). The postdiction superiority effect in metacomprehension of text. *Memory & Cognition, 29*, 62–67. https://doi.org/10.3758/BF03195741

Pilegard, C., & Mayer, R. E. (2015). Adding judgments of understanding to the metacognitive toolbox. *Learning and Individual Differences, 41*, 62–72. https://doi.org/10.1016/j.lindif.2015.07.002

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*, 459–470. https://doi.org/10.1016/S0883-0355(99)00015-4

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. The University of Michigan.

Pokay, P., & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology, 82*(1), 41–50. https://doi.org/10.1037/0022-0663.82.1.41

Rakoczy, K., Buff, A., & Lipowsky, F. (2005). Befragungsinstrumente [Assessment instruments]. In E. Klieme, C. Pauli, & K. Reusse (Eds.)*, 13. Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis". Materialien zur Bildungsforschung*. Deutsches Institut für Internationale Pädagogische Forschung. doi: URN: urn:nbn:de:0111-opus-31060.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Roth, A., Ogrin, S., & Schmitz, B. (2015). Assessing self-regulated learning in higher education: A systematic literature review of self-report instruments. *Educational Assessment, Evaluation and Accountability, 28*(3), 225–250. https://doi.org/10.1007/s11092-015-9229-2

Rovers, S. F. E., Clarebout, G., Savelberg, H. H. C. M., de Bruin, A. B. H., & van Merriënboer, J. J. G. (2019). Granularity matters: Comparing different ways of measuring self-regulated learning. *Metacognition and Learning, 14*(1), 1–19. https://doi.org/10.1007/s11409-019-09188-6

Rudolph, J., Niepel, C., Greiff, S., Goldhammer, F., & Kröner, S. (2017). Metacognitive confidence judgments and their link to complex problem solving. *Intelligence, 63*, 1–8. https://doi.org/10.1016/j.intell.2017.04.005

Saraç, S., & Karakelle, S. (2017). On-line and off-line assessment of metacognition. *International Electronic Journal of Elementary Education, 4*(2), 301–315.

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

Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemporary Educational Psychology, 19*, 143–154. https://doi.org/10.1006/ceps.1994.1013

Schraw, G. (1997). The effect of generalized metacognitive knowledge on test performance and confidence judgments. *The Journal of Experimental Education, 65*, 135–146. https://doi.org/10.1080/00220973.1997.9943788

Schraw, G. (2009a). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning, 4*, 33–45. https://doi.org/10.1007/s11409-008-9031-3

Schraw, G. (2009b). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415–429). Routledge.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460–475. https://doi.org/10.1006/ceps.1994.1033

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

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review, 7*(4), 351–371. https://doi.org/10.1007/BF02212307

Sobocinski, M., Järvelä, S., Malmberg, J., Dindar, M., Isosalo, A., & Noponen, K. (2020). How does monitoring set the stage for adaptive regulation or maladaptive behavior in collaborative learning? *Metacognition and Learning, 15*(2), 99–127. https://doi.org/10.1007/s11409-020-09224-w

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and self-regulated leaming constructs. *Educational Research and Evaluation, 10*, 117–139. https://doi.org/10.1076/edre.10.2.117.27905

Steuer, G., Engelschalk, T., Eckerlein, N., & Dresel, M. (2019). Assessment and relationships of conditional motivational regulation strategy knowledge as an aspect of undergraduates' self-regulated learning competencies. *Zeitschrift für Pädagogische Psychologie, 33*, 95–104. https://doi.org/10.1024/1010-0652/a000237

Tarricone, P. (2011). *The taxonomy of metacognition*. Psychology Press.

Vanderstoep, S., Pintrich, P., & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. *Contemporary Educational Psychology, 21*, 345–362. https://doi.org/10.1006/ceps.1996.0026

Vosniadou, S., Darmawan, I., Lawson, M. J., Van Deur, P., Jeffries, D., & Wyra, M. (2021). Beliefs about the self-regulation of learning predict cognitive and metacognitive strategies and academic performance in pre-service teachers. *Metacognition and Learning, 16*(3), 523–554. https://doi.org/10.1007/s11409-020-09258-0

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im studium. Ergebnisse zur faktorenstruktur und Reliabilität eines neuen fragebogens [Learning strategies of university students: Factor structure and reliability of a new questionnaire]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15*, 185–200.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45*, 267–276. https://doi.org/10.1080/00461520.2010.517150

Winne, P. H., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. Schunk, & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Taylor & Francis.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.

Winne, P. H., & Muis, K. R. (2011). Statistical estimates of learners' judgments about knowledge in calibration of achievement. *Metacognition Learning, 6*(2), 179–193. https://doi.org/10.1007/s11409-011-9074-8

Wirth, J., & Leutner, D. (2008). Self-regulated learning as a competence: Implications of theoretical models for assessment methods. *Zeitschrift für Psychologie, 216*, 102–110. https://doi.org/10.1027/0044-3409.216.2.102

Yuan, K.-H., & Bentler, P. M. (2004). On chi-square difference and z tests in mean and covariance structure analysis when the base model is misspecified. *Educational and Psychological Measurement, 64*, 737–757. https://doi.org/10.1177/0013164404264853

Zepeda, C. D., & Nokes-Malach, T. J. (2021). Metacognitive study strategies in a college course and their relation to exam performance. *Memory & Cognition, 49*, 480–497. https://doi.org/10.3758/s13421-020-01106-5

Zimmerman, B. J. (1998). Academic studying and the development of personal skill: A self-regulatory perspective. *Educational Psychologist, 33*, 76–86. https://doi.org/10.1207/s15326985ep3302&3_3

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166–183. https://doi.org/10.3102/0002831207312909

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676