Supporting Effective Self-Regulated Learning: The Critical Role of Monitoring

2

Thomas D. Griffin, Jennifer Wiley, and Carlos R. Salas

Abstract

This chapter explicates an empirically grounded and detailed theoretical framework for understanding the various components of self-regulated learning. A key distinction is articulated between metacognitive knowledge and metacognitive monitoring. It is argued that it is the accurate monitoring of learning experiences that is critical for effective self-regulation during learning, and that various accuracy measures for judgments of learning differ in how well they assess this construct of monitoring accuracy. Particular emphasis is placed on the importance of improving the relative accuracy of metacognitive monitoring skills, and that typical instruction in study strategies may not be sufficient to improve monitoring. The results of studies and manipulations that have resulted in superior monitoring accuracy are reviewed, and the implications for the development of learning technologies are discussed. A key observation is that in order to provide the opportunity for the development of effective regulatory skills, learning environments need to be careful not to deprive students of the opportunity to engage in self-regulation or monitoring of their own understanding.

Supporting Effective Self-Regulated Learning: The Critical Role of Monitoring

Imagine that a student has several homework assignments to complete in one night including reading several passages for biology and a set of readings in social studies. The readings for biology are on vision, taste, and the auditory system. The readings in social studies are a textbook passage on taxation without representation and two essays about the Boston Tea Party, one from an American perspective and one from a British perspective. For most daily schoolwork, students find themselves in situations such as this where they must regulate and monitor their own study behaviors. They must make important decisions such as when to read, what to read, how to read, and how much to read. Critical to this process is the ability to discriminate which readings

T.D. Griffin • J. Wiley (✉) • C.R. Salas
University of Illinois at Chicago, Chicago, IL, USA
e-mail: jwiley@uic.edu

R. Azevedo and V. Aleven (eds.), *International Handbook of Metacognition and Learning Technologies*, Springer International Handbooks of Education 28, DOI 10.1007/978-1-4419-5546-3_2,

have been understood well and which have not. This requires readers to actively and consciously monitor their ongoing learning progress in order to compare to a goal state or in relation to their progress on other tasks competing for their limited time and resources. Only with accurate monitoring will a student engage in effective self-regulated learning (SRL). Given the importance of monitoring to SRL, it is of great interest to find contexts that may improve the monitoring skills of students as they learn from text. However, pursuit of this goal requires clarity about what exactly the phenomenon of monitoring is and how it relates to the other components of SRL.

Monitoring and Its Place in SRL

Within literature on SRL, there are researchers who use similar terminology to refer to different constructs. This creates some confusion and potentially leads to incorrect inferences about what factors, individual differences, contexts, manipulations, and interventions influence particular aspects of SRL. Hacker (1998) commented on two different approaches to "monitoring." One approach, used primarily by cognitive psychologists, focuses upon learners' monitoring of ongoing learning via having students make overt judgments of their current level of understanding (judgments of learning, or JOLs) and comparing these to objective measures of the quality of their mental representations. The correspondence between these subjective and objective measures of learning is referred to monitoring accuracy or, more specifically, metamemory accuracy when the learning goal is memory (usually of word-pairs) and metacomprehension accuracy when the learning goal is comprehension (usually of texts). The focus of this research approach is to determine which conditions support accurate monitoring.

In contrast, another approach, used primarily by educational researchers, tends to use terms like "comprehension monitoring" more broadly to incorporate several kinds of monitoring, such as monitoring of goals, use and monitoring of strategies, as well as monitoring of learning. This approach generally attempts to improve SRL by supporting the use of particular learning and study strategies and utilizes assessments such as self-report scales of strategy knowledge and use, rather than focusing on accurate monitoring of ongoing learning.

The basis for both of these approaches to SRL was present in the original notion of metacognition put forth by Flavell (1979) 30 years ago, and both foci are still reflected in modern models of SRL. Flavell's original construct of metacognition was defined as "one's knowledge of one's own cognitive processes and products or anything related to them." The key components from Flavell's original theory of metacognition are depicted in Fig. 2.1. In Flavell's framework, metacognitive processes are designed to optimize one's cognitive *actions* in pursuit of learning *goals*. There are two major factors that determine the coordination of actions and goals. The first is the application of preexisting metacognitive *knowledge* about particular tasks, strategies, or a learner's abilities that can be used to select cognitive actions to increase learning. The second are metacognitive reactions to *experiences* of subjective internal states that occur as a result of the cognitive actions one executes and that reflect how learning is progressing. Metacognitive knowledge and experiences are distinct. Knowledge influences actions that in turn impact learning outcomes and can produce subjective experiences. However, as depicted by the recursive loop in Fig. 2.1, it is the internal metacognitive *experiences* associated with current attempts to learn that learners must monitor in order to judge their actual learning progress and make online revisions to their cognitive actions (i.e., regulation). Otherwise they will be guided only by incomplete and often erroneous prior knowledge. Later theories articulated that this experience monitoring process occurs at another level of awareness, the "meta" level of processing, because the subjective experiences that are being reflected on are the result of the cognitive processes or actions that students engage in at the "object" level (Fischer & Mandl, 1984; Nelson & Narens, 1990; for a recent discussion, see Griffin, Wiley, & Thiede, 2008). Consistent with Flavell, this model depicts monitoring as the

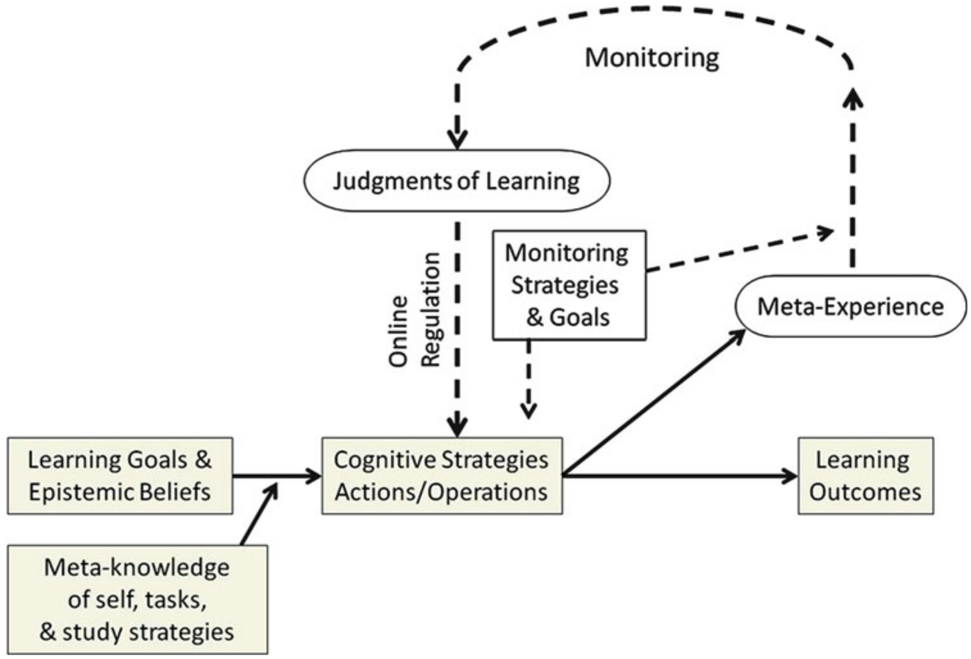

Fig. 2.1 Components of the self-regulated learning process

processing of one's own ongoing cognitive states (i.e., experiences), and regulation as the outcome of that processing whereby self-assessments of learning progress are used to alter the lower-level cognitive processing. Meanwhile, the implementation of strategies intended to improve learning occurs on the "object" level, as it represents a direct cognitive action. So while metacognitive knowledge contributes to cognitive processing, only monitoring of the ongoing learning experience has the quality of processing information about cognitive processes that defines *meta*cognitive processing. Figure 2.1 depicts knowledge states and processes that do not necessarily entail meta-level processing as the solid lines and boxes. The dashed lines and ovals entail meta-level processing where the learner is processing information about their own cognitive states.

Similarly, Flavell (1979) points out a critical distinction between *cognitive strategies* that are used to increase learning versus *metacognitive strategies* that are deliberately used to produce experiences that can be monitored to self-assess learning progress. Metacognitive strategies are essentially self-tests to evaluate learning. "Cognitive strategies are invoked to *make* cognitive progress, metacognitive strategies to *monitor* it" (Flavell, p. 909). Sometimes the entire distinction between cognitive and metacognitive strategy use rests in the learner's intended purpose for using a strategy. The same activity (e.g., asking oneself questions at the end of a chapter) could be employed as either type of strategy. If it is employed to deepen learning, it is a cognitive strategy. But if it is employed so that the learner can monitor and pay deliberate attention to the resulting subjective experiences to assess learning progress (such as the ease with which they answered the various questions), then it is a metacognitive strategy.

The fact that certain strategies direct learners to attend to the meta-level experiences resulting from self-testing actions is depicted in the center of the model as a moderating influence on the action-experience relationship. We refer to these as *monitoring strategies* to highlight the direct monitoring role served by only a subset of strategies, distinct from the object-level cognitive processing role of most strategies explored in SRL research. Only these experience monitoring strategies are part of the regulatory loop.

Influences of Metacognitive Knowledge on SRL

A common approach to improving SRL is to focus upon learners' awareness and use of study strategies. In a review of 201 empirical studies published from 2003 to 2007 in the major education journals on metacognition and SRL, Dinsmore, Alexander, and Loughlin (2008) reported that definitions of both constructs typically employ terms like monitoring and control. However, few of the studies actually assessed monitoring accuracy. Instead, most studies investigated awareness or use of study strategies, usually assessed with self-report measures.

Commonly used self-report inventories of metacognitive knowledge (Mokhtari & Reichard, 2002; Pintrich, Wolters, & Baxter, 2000), including the Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994) are dominated by items that assess general study strategies ("I summarize," "I read instructions carefully," and "I try to break studying down into smaller steps"), general self-beliefs ("I am good at remembering information"), and beliefs about contexts that impact learning ("I learn more when I am interested in the topic"). Learners receive higher scores when they report always using the same normatively preferred strategy, which means they are not actually regulating strategy use to specific contexts. Most of these scales have either no items (Moore, Zabrucky, & Commander, 1997) or very few items within a much larger scale (Mokhtari & Reichard, 2002; Pintrich et al., 2000) that explicitly assess monitoring strategies and goals. The MAI is somewhat of an exception with a few items designed to assess monitoring strategies, such as "I ask myself questions about how well I am doing while I am learning something new" and "I ask myself how well I've accomplished my goals." However, these items are typically analyzed as part of larger subscales that tap general information processing and study strategies, such as "I periodically review to help me understand important relationships" and "I summarize what I've learned." In addition, these subscales are typically analyzed as components of even broader latent constructs such as "Regulation of Cognition" which are combinations of many things including pre-task planning strategies, such as "I set specific goals before I begin a task" (Schraw & Dennison, 1994).

Some research has found that learners who score higher on these instruments do show superior text comprehension (Schraw & Dennison, 1994). Also, direct instruction in strategic reading has been shown to produce both changes in responses to these strategy inventories and improved comprehension or learning outcomes (Caverly, Nicholson, & Radcliffe, 2004; Pressley, 2002; Zimmerman, 2002). However, some critics have questioned whether these strategy inventories reflect actual strategy use since these self-reports have not been verified against converging measures of actual learning behaviors (Cromley & Azevedo, 2006).

The bottom pathway from left to right in Fig. 2.1 represents the direct influence that metacognitive knowledge can have on learning outcomes by impacting initial strategy selection during planning. This can entail generally effective strategies such as "summarize after reading" or context-dependent beliefs like "I learn more easily when interested" that interact with other a priori factors such knowledge about the task, topic, context, and beliefs about learning to determine strategy selection. In this model, metacognitive knowledge acts as an object-level cognitive process that directly impacts learning. This means that any observed relation between strategy use and learning outcomes can occur completely outside the regulatory loop, and the presence of such a relation cannot be used to determine whether monitoring is accurate or even if experiences are being monitored.

In fact, strong a priori commitment to a strategy that is generally effective could yield above average learning gains while also undermining online monitoring and regulation in certain situations where the strategy is suboptimal, resulting in inefficiency and costly use of resources. If strategy selection is based purely on a priori information, there is no opportunity for accurate monitoring of ongoing learning to play a role in SRL. Further, monitoring of actual learning outcomes in relation to strategy use is a critical source of information for updating and revising strategy knowledge in order to improve the efficacy of strategy choice on future learning trials. The exclusive reliance on a priori metacognitive knowledge will stagnate the long-term development of more accurate strategy knowledge because feedback from monitored learning outcomes is the presumed primary means by which errors in strategy knowledge are revised (Flavell, 1979; Winne & Hadwin, 1998).

The fact that many forms of strategy knowledge can have a direct effect on learning or an effect on strategy choice only in the planning stage is at least implicit in most models of SRL (e.g., Hacker, 1998; Nelson & Narens, 1990; Pintrich et al., 2000). For example, despite the highly recursive nature of Winne and Hadwin's (1998, Figure 12.1) SRL model, the arrows of influence show that preexisting knowledge of tasks and strategies can influence operations, cognitive products, and then performance without engaging the "monitoring" and "control" components at the heart of the model. This simply means that even though learners may have some awareness of task demands and may match this to a known strategy, once this initial plan is implemented, its influence on outcomes, products, and comprehension can occur without any online experience monitoring or responsive regulation.

Thus, a key point is that strategies that directly improve learning may or may not evoke metacognitive experiences that are useful for accurate monitoring. Students who are aware of these strategies or who report using them may be more effective learners, but these results will necessarily be unable to address whether they are better at online monitoring or regulation of their learning. Accurate online monitoring is important not only for revising strategies that have failed but also simply for knowing when the strategy needs to be repeated, due perhaps to idiosyncratic influences such as a brief distraction that limited its benefits. A priori strategy selection does not allow the learner to adapt to the numerous idiosyncratic contextual factors that foster and hinder comprehension processes as they actually occur. Judgments of learning that are based only in prelearning assumptions are not truly judgments of learning and cannot be used to modify and improve the initial strategies selected based upon those same assumptions.

Influences of Epistemic Beliefs on SRL

Similar issues can be raised about the burgeoning literature on learners' *epistemic beliefs* about the nature of knowledge and the process of knowing (with respect to the certainty, complexity, source, and potential revision of 'true justified' knowledge) and its influence on SRL (e.g., see Hofer & Sinatra, 2010). Research on *epistemic metacognition* has been shaped by models that construe epistemic beliefs as a type of general and abstract metacognitive knowledge (e.g., Hofer, 2004; Kitchener, 1983; Kuhn, 1999). However, rather than integrating epistemic beliefs into traditional models of SRL, this literature has largely attempted to construct a parallel model that repurposes monitoring as being in the service of "monitoring what [one] believe[s] to be true" and "monitoring and judging epistemic claims" for their truth status (Hofer, 2004, pp. 48–49) rather than monitoring of learning progress. Similarly, evaluative strategies are said to be regulated, such as by checking for internal logical inconsistencies in order to evaluate an argument's validity (Richter & Schmid, 2010) and generally increasing or decreasing one's efforts in evaluating a claim's veracity (Hofer, 2004).

In contrast to these parallel models of epistemic metacognition, Winne and Hadwin (1998) integrate epistemic beliefs into their more traditional SRL model as a component of metacognitive knowledge, where these beliefs serve as *cognitive conditions* that can foster use of certain learning strategies. For example, a belief that true knowledge is acquired effortlessly may promote the use of less effortful strategies. This expands upon and attaches an epistemic label to several kinds of general and abstract beliefs that Flavell (1979) also incorporated into the original model as part of metacognitive knowledge. Due to their level of abstraction and generality, epistemic beliefs might best be construed as determinants of learning goals that interact with knowledge of particular strategies to determine actual strategy use and cognitive actions.

The model in Fig. 2.1 depicts the effects of epistemic beliefs on learning as occurring via initial strategy selection without any impact on the metacognitive monitoring loop. Consistent with this suggestion, several recent studies have shown effects of epistemic beliefs on both the initial selection of more effective strategies (Bromme, Pieschl, & Stahl, 2010; Stahl, Pieschl, & Bromme, 2006) and on learning outcomes (e.g., Mason, Boldrin, & Ariasi, 2010; Muis & Franco, 2010). As with the study strategy literature, audiences might be misled to infer effects of epistemic beliefs on comprehension monitoring by the confusing use of traditional metacognitive terms employed in discussions of these findings. For example, Stahl and colleagues (2006) have described the effects of epistemic beliefs on learners' importance ratings of certain strategies for certain tasks in terms of superior *monitoring* and *calibration*. Mason et al. (2010, p. 85) have inferred superior *self-regulation* from pre-task self-reported general strategies. Muis (2008) has discussed *monitoring* effects in reference to engaging in behaviors that appear to reflect monitoring attempts, but these were not analyzed separately from non-monitoring behaviors. Most of these outcome measures do not reflect attempts to monitor ongoing learning progress or regulation in response to monitoring, and none reflect the accuracy of learners' monitoring. The potential effect of epistemic beliefs on actual comprehension monitoring and online regulation is still awaiting empirical confirmation.

In summary, this model highlights how metacognitive knowledge of context, goals, beliefs, and study strategies can influence learning and even regulation at the planning and selection stages without impacting monitoring of ongoing learning via reflection on experiences. Metacognitive knowledge serves to inform learners *what* strategies they should employ, but it is separate from the metacognitive processing that involves online monitoring of experience which can inform a learner *when* strategies are effective and when they need to be regulated, reapplied, or revised.

Experience Monitoring and Metacomprehension Accuracy

As posed above, a central element in models of SRL is the self-regulatory loop – the part of the model where a reader reflects on their own processing and alters their learning or study behaviors as a result. The regulatory loop depends on self-evaluation or judgments of learning (JOLs). Self-evaluation judgments, in turn, rely on cues. The quality of self-evaluation judgments depends largely on the quality of the cues that are used for the basis of these judgments. Such reasoning has been unpacked most extensively by Koriat in his cue-utilization theory. Koriat (1997) has discussed two classes of cues that learners use to draw inferences about their learning and future performance. One set are cues that are tied to the learner's internal online subjective experiences that reflect their cognitive processing in the specific situation. Because Koriat has been mainly concerned with judgments of learning during memorization tasks, he calls these mnemonic cues. These cues include the subjective sense of ease or fluency during learning (Benjamin & Bjork, 1996; Dunlosky & Nelson, 1992).

The other kinds of cues are tied to objective features of the learning situation that are either *intrinsic* to the materials and task demands (e.g., relatedness of word-pairs, memory of details versus conceptual application) or *extrinsic* to the task or stimuli, but instead related to the context (e.g., how many times items were studied or what strategy was used). These knowledge-based cues bypass the monitoring of subjective experience. Instead, people may make judgments based on their perceptions of the general effectiveness of certain strategies.

Although the cue-utilization theory was developed with reference to metamemory monitoring of rather simple materials like word-pairs, it can be adapted to metacomprehension of complex texts. Such an adaption is reflected in the model by Griffin, Jee, and Wiley (2009) that distinguishes heuristic from representation-based cues that can be used for self-evaluation. Representation-based cues, like mnemonic cues, are tied to subjective online experiences that reflect processing during learning and the quality of the mental representation that a learner has actually formed. Heuristic cues are those based in a priori general assumptions about topic interest, domain knowledge, ability, and text and task features. The model proposes that these heuristic cues (which comprise metacognitive knowledge) may have modest predictive validity because they refer to things that can have some influence on learning, but they are insensitive to idiosyncrasies of the specific learning situation and can therefore be erroneous. For example, the heuristic knowledge that one is good at multiple-choice tests may predict higher than average overall performance on such tests, but will be of no help in predicting whether one will do better on a test about the Irish potato famine versus a test about earthquakes. In addition to capturing Koriat's key distinction between cue types and providing a reason why mnemonic cues are generally more valid, this heuristic-representation distinction maps rather clearly onto Flavell's (1979) knowledge-experience distinction. The subjective experiences that readers need to monitor are those that arise from processes of building a mental representation of the meaning of a text. However, when metacognitive knowledge is used as a heuristic that directly influences judgments of learning, then it bypasses the active monitoring process and use of representation cues. In essence, the judgments are no longer about learning, but are merely performance predictions based on a priori knowledge of factors that may or may not have some impact on whatever learning actually occurred.

Measures of Monitoring Accuracy

In the metacomprehension literature, a standard approach has been developed to assess the accuracy of these self-evaluations and the ability of students to monitor their ongoing comprehension processes (Glenberg & Epstein, 1985; Maki, 1998). In the typical metacomprehension paradigm, participants read a series of texts on a variety of topics, then rate their comprehension of each text, and complete a test for each text. Following the lead of metamemory research on paired associate learning (Nelson, 1984), a person's monitoring accuracy is operationalized as the intraindividual correlation between a person's comprehension ratings and actual test performances across the set of texts. More accurate self-evaluation or greater monitoring accuracy is indexed by stronger correlations. A standard term for this predictive accuracy measure is *relative metacomprehension accuracy*. This relative accuracy paradigm targets ongoing active monitoring of actual learning progress independent from either the level of progress itself (Nelson, 1984) or the learner's ability to make heuristic guesses about average progress based on general a priori beliefs about themselves or the task (Griffin et al., 2009).

In addition, this paradigm attempts to tap into the kinds of decisions a student must make as they decide among homework activities. If a student does not accurately differentiate well-learned material from less-learned material, time could be wasted studying material that is already well learned while no time would be devoted to material that has not yet been adequately learned. Students will also fail to realize when current study strategies are not working and new ones are needed. Consistent with this proposition, relative monitoring accuracy has been demonstrated to relate positively to self-regulated learning outcomes (Thiede, Anderson, & Therriault, 2003).

There are several reasons why relative accuracy has become the standard for determining monitoring accuracy. The other measures of metacognitive judgments (i.e., confidence bias, absolute accuracy) differ from relative accuracy in important ways. A central premise of research on metacognitive monitoring and SRL is the recognition that students do not have unlimited time to engage in study, and principled decisions need to be made about what should be studied or restudied for efficient self-regulated learning. Only measures of relative metacomprehension accuracy address this aspect of SRL.

Beyond this ecologically valid feature of the relative accuracy paradigm, a major reason for the increasing dominance of the relative accuracy paradigm in metacomprehension research is because it represents a measure of monitoring that is not heavily dependent upon average test performance (Nelson, 1984; Yates, 1990). Other methods used to assess monitoring accuracy simply compute the difference between judgments of learning and objective performance. As such, they are just as dependent upon how well a learner generally performs as they are on their skill in monitoring that performance. These methods include *absolute accuracy* which computes the unsigned absolute difference and *confidence bias* which computes the signed difference (Maki, 1998). Two people with poor monitoring skills who both just use a midpoint-of-the-scale heuristic can have drastically different absolute accuracy or confidence bias just because of differences in performance. Not only will the readers differ in accuracy despite no differences in their monitoring process, but one could have extremely high accuracy even though neither are actually monitoring at all and both are merely using a general anchoring heuristic. With relative accuracy, both readers would wind up with a poor accuracy score close to a correlation of zero, which would validly reflect the fact that they failed to monitor. In addition, since average performance levels often reflect relatively stable individual differences and can be systematically impacted by features of the learning context, all of these non-metacognitive factors will systematically produce differences in absolute accuracy and confidence bias, even when there are no differences in either the judgments themselves or the psychological processes that give rise to them. Differences on these measures may not reflect anything about metacognitive processes or skills.

Confidence bias brings even more interpretive problems, because it is not a linear measure of degree of accuracy but rather of the amount of directional bias in whatever errors exist. A score of zero reflects a lack of directional bias, and positive scores reflect more overconfidence errors while negative scores reflect more under-confidence errors (see Yates, 1990). One person with a higher score than another can be either less under-confident or more overconfident and either less accurate or more accurate depending on where each of these two people being compared happen to be in relation to the zero point. Group means for confidence bias reflect whether more people were over or under confident and do not represent the average level of accuracy of individuals. As a result, differences in this measure reflect neither monitoring nor accuracy.

One benefit of relative accuracy that has not been previously emphasized is that the independence of relative accuracy from average performance makes it the only measure of accuracy that necessarily reflects the actual monitoring of ongoing learning. Because it is not dependent upon average performance, high relative accuracy cannot be achieved by the use of heuristic meta-knowledge, even when that knowledge is accurate. Instead, high relative accuracy requires active attention to the ongoing learning process and its variable outcomes. Whether a reader has accurate knowledge about their own general skill in science learning might greatly impact both their absolute accuracy and their confidence bias, but this heuristic will be of little relevance in predicting their understanding of a text on volcanoes relative to their understanding of a text on evolution. This positive feature of relative accuracy makes it a superior measure of a students' ability to actually monitor ongoing learning processes which is the heart of the self-regulation processes in SRL. Absolute accuracy and confidence bias measures are not capable of discriminating real monitoring from either performance effects or the reliance on heuristic judgments that bypass monitoring processes in predicted overall performance levels.

As shown in Fig. 2.1, knowledge of and use of study strategies that determine what actions and operations are enacted play a very different role in SRL than experience monitoring. Study strategies are largely object-level constructs that guide actions which may or may not happen to evoke attention to meta-level experiences as represented by the regulatory loop. Yet, a number of researchers have used terminology such as metacognitive monitoring to refer to monitoring of one's strategy use. Even when increased knowledge of study strategies has positive effects on learning, it may not affect the regulation process. In fact, if learners become overly confident in their existing strategy-outcome beliefs, they may rely more heavily on these beliefs at the expense of monitoring subjective experience. Thus, theoretically, strategy instruction cannot be assumed to lead to better monitoring of learning progress and in fact could harm it.

In addition to a lack of theoretical basis to generally assume a positive effect of strategy knowledge or use on monitoring accuracy, there is a lack of empirical support. The few studies in the metacognitive knowledge literature that have measured JOLs have operationalized accuracy with the problematic measures of absolute accuracy or confidence bias (e.g., Schraw & Dennison, 1994), while failing to account for the non-monitoring influences of average performance and heuristic cues that plague these measures. Learning environments cannot be presumed to have improved monitoring processes unless improvements in JOL accuracy can be demonstrated independent from any effects on performance itself. And, unless a measure of relative accuracy is employed, then claims of benefits in monitoring skills are unwarranted.

Improving Monitoring Accuracy with a Valid Cues Approach

Both models and data suggest that accurate metacognitive monitoring of ongoing learning is central to effective regulation of study (e.g., Metcalfe, 2002; Nelson & Narens, 1990; Thiede, Griffin, Wiley, & Redford, 2009; Winne & Hadwin, 1998; Zimmerman, 2002). Because accurate monitoring is so critical for effective SRL, it is of great concern that the typical finding from the metacomprehension literature is that levels of monitoring accuracy are quite low. Several independent reviews have reported that the mean intraindividual correlation between comprehension ratings and test performance across numerous studies is only about +0.27 (Dunlosky & Lipko, 2007; Lin & Zabrucky, 1998; Maki, 1998). A recent comprehensive review of all published studies of relative monitoring accuracy for learning from text done in the last 30 years arrived at the same figure of 0.27 for the average among baseline conditions (Thiede et al., 2009). This review also showed that the majority of manipulations have little effect in improving this accuracy. The above analysis of cue validity suggests that in order to be accurate, students need to be monitoring cues directly related to reading experiences and not just relying on heuristic bases for their judgments. However, there are many levels on which one can attempt to monitor their reading processes, and only some of these are predictive of comprehension. When considering learning from text, we must bear in mind that a text can be processed at several levels from surface memory of the exact words to constructing a conceptual model of the meaning of the text (Graesser, Millis, & Zwaan, 1997; Kintsch, 1998). To make accurate judgments of comprehension, readers need to reflect specifically on experiences that correspond to the level of representation the learning task requires. Because it is a person's situation model that largely determines his or her performance on tests of comprehension (Kintsch, 1998; McNamara et al., 1996), metacomprehension monitoring will be most accurate when situation-model level cues are utilized (Rawson, Dunlosky, & Thiede, 2000; Wiley, Griffin, & Thiede, 2005). For example, Thiede, Griffin, Wiley, and Anderson (2010) observed that most readers self-report that they base their judgments of learning upon heuristic judgment cues related to text features (e.g., "the text was long") or upon beliefs about their own skill and familiarity with the topic. Readers' reported use of representation cues was largely limited to how much of the text they could remember. Both the reliance on heuristic and immediate memory cues were associated with poor monitoring accuracy, while those few readers who did self-report relying upon situation-model-level cues (like the ability to explain a causal process described in the text to someone else) tended to have superior monitoring accuracy. The assumption that monitoring accuracy can be improved by shifting readers to rely more upon valid situation-model cues is the foundation for our work described below.

Instantiating a Relative Accuracy Paradigm

In these studies, our goal has been attempting to find conditions that improve readers' ability to accurately judge their own level of comprehension from text using a standard relative accuracy paradigm. The texts are approximately 1,000 words long and are on science topics such as the vision system. Sets of five to six texts on different topics are generally used. Students read all texts, then they are asked to judge their level of comprehension for each text ("How many items do you think you will get correct on a 5 item test?"), and then they take comprehension tests in the same order as reading. The comprehension tests consist of five multiple-choice items tapping inferences that follow from each text.

Design Considerations for Texts

Wiley et al. (2005) pointed out that the design of the expository texts and comprehension tests are both critical to examining metacomprehension accuracy. Only texts that have clearly distinguishable surface and situation-model representations and only test questions that can be answered using just one or the other representation will lead to interpretable results. Thus, we use explanatory science texts for which the situation model is not entirely explicit within the surface model of the text. Since creating the situation model for a text involves generating inferential connections, it is important to construct texts that can test whether the reader is making connections beyond what is explicitly stated. Our own texts typically describe a complex causal relation (e.g., the relation between continental and ocean plates and the emergence of volcanoes). For example, a well-developed situation model for the volcano text would contain inferential links such as "the least likely place for a volcano is in the center of a plate." The key here is that this connection needs to be constructed by the reader. The text itself does not contain this statement. Based on previous research (Kintsch, 1998), we believe that comprehension is best represented by a person's situation model for a text, and the quality of reader's understanding of a text can best be discerned by assessing whether the person can recognize causal inferences implied by a text (Trabasso & Wiley, 2005; Wiley & Myers, 2003). When the test performance being predicted reflects the quality of a reader's situation model, then the accuracy of the monitoring judgments represents meta*comprehension,* as opposed to meta*memory* for explicitly stated idea units within a text. Although readers must also comprehend explicitly stated ideas, researchers must take care to create tests that require actual understanding of those ideas rather than mere memory for words.

Design Considerations for Tests

We also have specific considerations for the design of our comprehension tests. One important feature is that they contain more than one or two items. Weaver (1990) addressed the weaknesses associated with assessing comprehension monitoring with limited items per text. In particular, he argued that a one-item test does not provide a reliable measure of comprehension. Moreover, using a one-item test creates an issue of content coverage, where computed monitoring accuracy is highly contingent upon the arbitrary overlap between what portion of the text the test covers and what portion the readers emphasized in their judgment. Thus, it is important to use tests with multiple items that assess comprehension of the majority of the content presented in the text.

Perhaps more important, the tests must also provide a valid measure of comprehension (i.e., tap the situation model of the text). With these concerns in mind, we have developed multiple-choice tests (following Royer, Carlo, Dufrense, & Mestre, 1996 and Wiley & Voss, 1999) that directly tap understanding of text content by asking students to verify inferences that follow from the texts. Performance on the inference tests that we have developed reliably correlates with other learning assessments, including performance on "how" and "why" essay questions (Sanchez & Wiley, 2006; Wiley et al., 2009), as well as with performance on the Nelson Denny (Griffin et al., 2008).

Design Considerations for Judgments

The valid cues approach suggests that the more strongly a cue is diagnostic of the mental representation that will determine test performance, the more valid and predictive of performance it will be. An extreme illustration of this point is that postdictions are generally very accurate (Maki & Serra, 1992; Pierce & Smith, 2001). A postdiction is when a person simply predicts future performance based on a prior test that assesses the same mental representation. The cues that are generated by the initial test with the same items are directly diagnostic for later performance, which explains the postdiction superiority effect. However, note that providing learners with the actual test questions and the experience of answering them circumvents the need for engaging in monitoring of the learning experience. Postdiction judgments are more accurate because they do not rely upon the metacognitive system and do not require any actual metacognitive monitoring which is what learners struggle with. For this reason, predictive judgments are more useful as a measure of online monitoring processes. It is also useful if the judgments are made in the same metric as the test scores.

Supporting Access to Valid Comprehension Cues

In earlier work (Thiede et al., 2003), having students engage in delayed generation tasks (keyword listing or summaries) after reading produced unprecedented levels of metacomprehension accuracy compared to an immediate generation control group. Because both groups engaged in generation, an implication was that the delay itself was responsible. However, Thiede, Dunlosky, Griffin, and Wiley (2005) conducted a series of follow-up studies that independently manipulated delay and generation tasks. Simply delaying judgments did nothing to accuracy and neither did having readers perform non-generative tasks at a delay, such as reading a list of keywords or being prompted to "think about the text." The key to producing better monitoring accuracy was in making readers perform a specific type of generative self-test. In this case, these generation tasks (summary or keyword listing) only yielded benefits when performed at a delay. This is because these tasks can be done using surface memory when performed immediately, but the surface representation decays with a delay while the situation model is more robust over time (Kintsch, Welsch, Schmalhofer, & Zimny, 1990). It was not delaying judgments themselves but being directed to perform a delayed generation task as a self-test that increased readers' access to the appropriate representation cues and improved monitoring accuracy.

Griffin et al. (2008) provided further evidence that certain types of self-testing targeted toward situation-model cues can increase accuracy. One study employed self-explanation as the type of self-test designed to increase access to valid cues. Readers who engaged in a self-explanation task while reading had significantly higher metacomprehension accuracy than those who simply reread. Self-explanation requires readers to simultaneously construct and self-test their situation model by asking themselves how certain ideas fit together with the theme of the text (Chi, 2000; Wiley & Voss, 1999). Accuracy improved even without delaying judgments. Self-explanation directly involves the situation model, making the timing less relevant to what cues are accessed by it, unlike keyword lists and summaries that could be based largely in a surface representation when performed immediately. Another important aspect of this study was that there was actually no effect of self-explanation on test performance itself. One should not view the lack of learning gains in this study as conflicting with other research on self-explanation, since these students received neither training in how to self-explain nor did they have the opportunities for restudy that have supported better learning in other studies (Chi, 2000; McNamara, 2004). Instead, the lack of effects on performance allows for the conclusion that self-explanation had its effect on monitoring, since performance was not affected but accuracy was improved.

Another study reported by Griffin et al. (2008) has shown that simple rereading can improve metacomprehension accuracy, but only for readers with limited attentional resources or low comprehension skill. These effects were interpreted as demonstrating that readers with limited or taxed attentional resources during a single reading can use a second reading to attend to important online experience-based representation cues. Without the resources to attend to these cues during a first reading, readers are forced to rely more heavily on heuristic cues. Together these studies from the first phase of our research program suggest that the key factor in utilizing valid representation cues is having access to these cues, both by being able to attend to them when available and making them more available by employing self-tests designed to target the appropriate level of representation. This work utilizing delayed generation, rereading, and self-explanation has been successful at producing uncommon levels of monitoring accuracy, raising intraindividual correlations between judgments and performance from the usual 0.27 to above 0.6 in most cases.

Supporting the Selection of Valid Comprehension Cues

The interventions previously described direct readers to engage in cognitive actions designed to evoke certain metacognitive experiences and make valid representation cues more accessible. Although this increase in accessibility makes valid cue use more likely, optimal cue use will also require readers to actively discriminate and select among those cues available to them. If texts and tests require students to gain conceptual understanding, for example, of scientific processes and phenomena from expository text, then it is important to prompt students to override the "reading for memory" setting evident in their self-reported selection of memory cues over situation-model cues (Thiede et al., 2010). Readers need to realize that their goal for reading is to try to understand how or why a phenomena or process occurs and that the questions they will be asked will depend on making connections and causal inferences across sentences, in order to engage in monitoring of the most relevant experiences. This influence of cue selection on monitoring is depicted in Fig. 2.1 as the arrow from *monitoring goals* that intersects the link between *experiences* and *monitoring*. In terms of Winne and Hadwin (1998), we suggest that in order to engage in effective SRL, learners need meta-knowledge of standards on which their learning can be evaluated. In terms of the present model, learners' monitoring goals need to reflect the appropriate level of understanding or type of learning, so that they can selectively attend to and make use of those metacognitive experiences that reflect this level of understanding.

Thus, in a second series of experiments, we attempted to shape the selection of valid cues by influencing learners' test expectancies with an explicit statement about the inferential nature of the final test items they should expect and the need to make connections between different parts of a text. Readers were also given practice texts and tests with inference items to set the expectation. This manipulation has been highly effective in improving relative monitoring accuracy (Thiede, Wiley, & Griffin, 2011). In additional studies, we have found that when combining this test-expectancy manipulation with a self-explanation instruction, the two interventions had independent effects, suggesting that both cue accessibility and cue selection are determining accuracy and are distinct contributors to cue use (Wiley et al., 2008).

Negative Effects of Providing Feedback

It is critical to note that in the above test-expectancy studies, students were not provided with any performance feedback on the practice tests. The effects of test expectancy were assessed by a transfer paradigm in which monitoring goals had to be generalized from the practice trials and applied to new texts and tests.

Given that attention to internal experiences defines metacognitive monitoring, externally provided performance feedback during practice tests may short-circuit effective monitoring of ongoing learning by shifting readers' attention from internal to external cues. Overt judgments of learning will no longer be based in inferences derived from the experience monitoring process, but rather based in the externally provided information, such as simply anchoring all future judgments on the numerical score one received on the previous tests. When external feedback is predictive of future performance, such as when the future tests are on the same material, the accuracy of JOLs may increase even though readers are no longer truly monitoring. However, when accurate JOLs depend upon actual meta-level monitoring because the feedback on past performance is not related to future performance, then JOL accuracy could be harmed by feedback.

An example of such a scenario is when learners' might receive feedback on their performance on one set of texts, but later need to monitor their learning for a new set of texts on different topics. On the one hand, the practice tests provide a basis for abstracting a transferable expectancy they can use to guide their monitoring during future texts, but on the other hand the concrete numerical performance scores on the practice tests may become the basis for future judgments on other texts, without regard to the fact that they are about different topics and thus require an independent judgment. In other words, the readers might merely transfer the concrete numerical performance scores from one text to another rather than the more abstract concept about the general nature of the type of test and level of comprehension required.

We tested this scenario by employing the same test-expectancy paradigm previously described, but added two feedback conditions. Both feedback conditions were identical to the inference test-expectancy condition, except they also gave readers performance feedback (i.e., number of questions answered correctly) for the practice inference tests. One of the feedback conditions also reminded readers of their JOLs in relation to their actual practice performance. If feedback undermines experience monitoring, then the benefits of having a valid monitoring goal created by inference test expectancy should disappear when that expectancy is accompanied by prior performance feedback. The results supported this hypothesis, revealing that the notably improved monitoring accuracy by providing inference test expectancies ($r=0.49$) versus control ($r=0.15$) was completely eliminated by simply adding feedback on practice test performance ($r=0.21$). Apparently, readers focused upon the external concrete practice feedback and failed to transfer an expectancy about the more general nature of the tests. We do not know whether the participants in this study actually failed to engage in monitoring due to the feedback or whether they simply failed to use the cues derived from that monitoring when making their judgments. But, it is clear that readers were unduly influenced by their past performance scores when predicting future performance, even though those scores had little relevance. Obviously, feedback can have a number of positive effects on learning. The point here is that the development of accurate monitoring skills may be best aided by practice tests that are not accompanied by concrete numerical performance feedback.

Implications for the Design of Learning Technologies

This chapter has attempted to explicate an empirically grounded and detailed theoretical framework for understanding the various related but distinct components of SRL. The emphasis has been upon the importance of accurate metacognitive monitoring for engaging in effective regulation of learning. Understanding these conceptual and theoretical issues is critical for those who seek to develop instructional environments to foster the development of self-regulation skills. In particular, we highlight a few observations about the implications of this approach for the design of learning environments.

Regulation Is a Process of Making Decisions

Effective self-regulated learning involves decisions about what to read next, what to reread, and what strategies to apply as you are reading. If you take those decisions away from the learner, then you rob them of the opportunity to develop skills in regulating their learning. Learning environments may not be able to support both the most efficient learning at the cognitive level and the development of regulatory skills at the metacognitive level simultaneously. Conditions that aid learning of content (such as by matching difficulty of the learning task to each student's ability or prescribing strategy use) may lead to improvements in learning for that unit when the student is supported by the system. However, they may obviate the need for the student to grapple with difficulties and make their own choices about what to study next and how to study it, which may have negative consequences for their ability to engage in effective SRL in new, unsupported contexts.

Regulation Is a Process of Self-Evaluation

If you give feedback, then readers no longer need to self-evaluate. As we have shown above, giving feedback can be problematic for monitoring accuracy. Dictating the use of a particular learning strategy also obviates the need for self-evaluation. To support SRL, learning environments need to support self-testing and online monitoring strategies. Theoretically, the only types of strategy knowledge and use that should directly impact monitoring are those that explicitly direct learners' attention toward metacognitive processing, such as attending to the ease with which one can summarize information or answer self-generated questions as an indicator of comprehension. However, these *metacognitive* monitoring strategies are not well represented on the most commonly used inventories. They also do not seem to be the type of strategies that are taught or supported in most learning technology environments. Indeed many intelligent tutoring and cognitive tutoring systems remove the need to monitor one's own level of performance and regulate actions as the learning technology is often designed to monitor students' learning for them.

More often, the strategies that are supported by learning technologies are study strategies that more directly support learning. Learning environments designed to foster students' knowledge of effective study strategies should avoid breeding excessive confidence in the global efficacy of specific strategies. Rather, students could be taught a repertoire of strategies, made aware that strategy effectiveness is context dependent, and prompted to always monitor their learning progress and reassess effectiveness of each strategy in each particular learning context. This decision process would help to support reflection and regulation skills, especially if coupled with instruction in strategy use explicitly for the purpose of monitoring, such as self-testing or self-explanation.

Final Thoughts

Accurately monitoring one's current state of understanding during a cognitive task is a central feature of effective control and self-regulation that impacts learning for both that task and potentially for future tasks. Monitoring one's "experiences of puzzlement or failure," such as a "sense that you do not yet know a certain chapter in your text well enough," is critical for creating new subgoals, applying alternate strategies, and revising one's metacognitive knowledge about the effectiveness of the strategies (Flavell, p. 908). In other words, the monitoring of the dynamic and changing states of one's learning progress is what tells a reader *when* they need to intensify, reduce, stop, or alter the cognitive learning strategies being employed and is what informs the learner what strategies should be modified, deleted, or added to the strategy knowledge base for use on future tasks. Without this monitoring of actual learning, a learner is not engaging the heart of self-regulated learning. Further, without studies that directly assess the accuracy of this monitoring, it is difficult to draw conclusions about which learning technologies may improve the monitoring skills needed for effective SRL. To provide the opportunity for the development of effective regulatory skills, learning environments need to be careful not to deprive students of the opportunity to monitor their own understanding.

References

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 309–38). Hillsdale, NJ: Erlbaum.

Bromme, R., Pieschl, S., & Stahl, E. (2010). Epistemological beliefs are standards for adaptive learning. *Metacognition and Learning, 5*, 7–26.

Caverly, D. G., Nicholson, S. A., & Radcliffe, R. (2004). The effectiveness of strategic reading instruction for college developmental readers. *Journal of College Reading and Learning, 35*, 25–46.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Hillsdale, NJ: Erlbaum.

Cromley, J. G., & Azevedo, R. (2006). Self-report of reading comprehension strategies: What are we measuring? *Metacognition and Learning, 1*, 229–247.

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*, 429–444.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science, 16*, 228–232.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20*, 374–380.

Fischer, P. M., & Mandl, H. (1984). Learner, text variables, and the control of text comprehension and recall. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 213–254). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*, 906–911.

Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 702–718.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163–189.

Griffin, T. D., Jee, B. D., & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition, 37*, 1001–13.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition, 36*, 93–103.

Hacker, D. J. (1998). Self-regulated comprehension during normal reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 165–191). Mahwah, NJ: Erlbaum.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: thinking aloud during online searching. *Educational Psychologist, 39*, 43–56.

Hofer, B. K., & Sinatra, G. M. (2010). Epistemology, metacognition, and self-regulation: musings on an emerging field. *Metacognition and Learning, 5*, 113–120.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29*, 133–159.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: a three-level model of cognitive processing. *Human Development, 26*, 106–116.

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.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Research, 28*, 16–26.

Lin, L., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*, 345–391.

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). Hillsdale, NJ: Erlbaum.

Maki, R. H., & Serra, M. (1992). Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology, 84*, 200–210.

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: evaluating and learning online information. *Metacognition and Learning, 5*, 67–90.

McNamara, D. S. (2004). Self-explanation reading training. *Discourse Processes, 38*, 1–30.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? *Cognition and Instruction, 14*, 1–43.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology. General, 131*, 349–363.

Mokhtari, K., & Reichard, C. A. (2002). Assessing student's metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*, 249–259.

Moore, D., Zabrucky, K., & Commander, N. E. (1997). Validation of the metacomprehension scale. *Contemporary Educational Psychology, 22*, 457–471.

Muis, K. R. (2008). Epistemic profiles and self-regulated learning. *Contemporary Educational Psychology, 33*, 177–208.

Muis, K. R., & Franco, G. (2010). Epistemic profiles and metacognition: support for the consistency hypothesis. *Metacognition and Learning, 5*, 27–45.

Nelson, T. O. (1984). A comparison of current measures of feeling-of-knowing accuracy. *Psychological Bulletin, 95*, 109–133.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–141). New York: Academic.

Pierce, B. H., & Smith, S. M. (2001). The postdiction superiority effect in metacomprehension of text. *Memory & Cognition, 29*, 62–67.

Pintrich, P. R., Wolters, C., & Baxter, G. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 43–97). Lincoln, NE: Buros Institute of Mental Measurement.

Pressley, M. (2002). Metacognition and self-regulated comprehension. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 184–200). Newark, De: International Reading Association.

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition, 28*, 1004–1010.

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self- regulated learning. *Metacognition and Learning, 5*, 47–65.

Royer, J. M., Carlo, M. S., Dufrense, R., & Mestre, J. (1996). The assessment of levels of domain expertise while reading. *Cognition and Instruction, 14*, 373–408.

Sanchez, C., & Wiley, J. (2006). Effects of working memory capacity on learning from illustrated text. *Memory & Cognition, 34*, 344–355.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460–475.

Stahl, E., Pieschl, S., & Bromme, R. (2006). Task complexity, epistemological beliefs, and metacognitive calibration: An exploratory study. *Journal of Educational Computing Research, 35*, 319–338.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*, 66–73.

Thiede, K. W., Dunlosky, J., Griffin, T. D., & Wiley, J. (2005). Understanding the delayed keyword effect on metacomprehension accuracy. *Journal of Experiment Psychology: Learning, Memory and Cognition, 31*, 1267–1280.

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes, 47*, 331–362.

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition and self-regulated learning* (pp. 85–106). Mahwah, NJ: Erlbaum.

Thiede, K. W., Wiley, J., & Griffin, T. D. (2011). Test expectancy affects metacomprehension accuracy. *British Journal of Educational Psychology, 81*, 264–273.

Trabasso, T., & Wiley, J. (2005). Goal plans of action and inferences during comprehension of narratives. *Discourse Processes, 39*, 129–164.

Weaver, C. A. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 214–222.

Wiley, J., Goldman, S., Graesser, A., Sanchez, C. A., Ash, I. K., & Hemmerich, J. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal, 46*, 1060–1106.

Wiley, J., Griffin, T., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology, 132*, 408–428.

Wiley, J., Griffin, T. D., & Thiede, K. W. (2008). To understand your understanding you must understand what understanding means. In V. Sloutsky, B. Love, & K. McRae (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society*. Cognitive Science Society: Austin, TX.

Wiley, J., & Myers, J. L. (2003). Availability and accessibility of information and causal inferences from scientific text. *Discourse Processes, 36*, 109–129.

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources. *Journal of Educational Psychology, 91*, 301–311.

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: Erlbaum.

Yates, J. F. (1990). *Judgment and decision making*. Englewood Cliffs, NJ: Prentice-Hall.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*, 64–72.