Concurrent and Task-Specific Self-Reports

Ivar Bråten

University of Oslo, Norway

Joseph P. Magliano

Georgia State University, USA

Ladislao Salmerón

University of Valencia, Spain

August 2019

Bråten, I., Magliano, J.P., & Salmerón, L. (2019). Concurrent and task-specific self-reports. In D.L. Dinsmore, L.K. Fryer, & M.M. Parkinson (Eds.), Handbook of strategies and strategic processing: Conceptualization, intervention, measurement, and analysis. New York: Routledge.

Introduction

In their landmark review of research on "learning, remembering, and understanding," Brown, Bransford, Ferrara, and Campione (1983) described a major metaphorical shift occurring during the late 1960s and early 1970s: a shift away from a passive learner responding to environmental influences toward an active learner using strategies in the service of acquiring, retaining, and understanding information. Although those authors acknowledged that it was not particularly clear what was strategic and what was not, their review of the progress made in this area of research during the 1970s and early 1980s focused on "deliberate plans and routines called into service for remembering, learning, or problem solving" (Brown et al., 1983, p. 85).

The defining attributes of strategies were, sometimes fiercely, debated in the following years, especially regarding the attribute of consciousness or intentionality (e.g., Paris, Newman, & Jacobs, 1985; Pressley, Forrest-Pressley, Elliot-Faust, & Miller, 1985). However, more recent views seem to have converged on the idea that strategies involve effortful, intentional, and planful processing (Afflerbach, Pearson, & Paris, 2008; Alexander, Graham, & Harris, 1998; Kendeou & O'Brien, 2018). In this way, strategies are distinguished from skills, which denote automatic, unintentional, and routinized information processing (Afflerbach et al., 2008). In accordance with this distinction, we define strategies as forms of procedural knowledge that individuals intentionally and planfully use for the purpose of acquiring, organizing, or elaborating information, as well as for reflecting upon and guiding their own learning, comprehension, or problem solving (cf., Alexander et al., 1998; Bråten & Samuelstuen, 2004; Weinstein, Husman, & Dierking, 2000). Thus, when individuals perceive a discrepancy between a desired outcome and their current state of learning, comprehension, or problem solving, and automatic skills cannot get them to the goal, they may decide to invest effort in strategic processing in order to reduce or eliminate that discrepancy (Alexander et al., 1998).

Further, an important distinction has concerned superficial versus deeper processing strategies, with level of processing related to the extent to which information is reorganized or transformed during learning, comprehension, or problem solving. This distinction is consistent with Bereiter and Scardamalia's (1987) classic distinction between knowledge-telling and knowledge-transforming approaches, with the former involving a superficial engagement suitable for reproduction of information in the same or similar form and the latter involving an active transformation of information suitable for generating new meaning or insights. With reference to our definition of strategies, individuals may use superficial strategies, such as selection and rehearsal, to acquire new information, and deeper level strategies, such as constructing summaries and drawing inferences, to organize and elaborate information. The latter part of our definition, referring to reflection and self-guidance, captures metacognitive strategies such as planning, monitoring, control, and evaluation of processing as well as performance (Veenman, 2016).

Needless to say, valid measurement of strategic processing is essential, for example, to understand how the different components or aspects of strategic processing work together with motivation and other forms of cognition, how they come into play and influence task completion and performance within and across domains, and how strategic processing develops over time, both naturally across the lifespan and as a consequence of deliberate strategies instruction. Given the effortful, conscious, and intentional nature of strategic processing, self-report methodologies, indeed, seem applicable in gauging such thinking. Still, there are several caveats concerning some of these methodologies (Tourangeau, 2000;

Veenman, 2011), requiring careful consideration of which forms of self-reports may produce results that can be trusted. In this chapter, we generally define self-reports as utterances or answers to prompts or questions provided by an individual about his or her cognitions and actions during learning, comprehension, or problem solving. We discuss concurrent and task-specific self-reports, in particular, focusing on the possibilities and challenges of using such approaches to measuring strategic processing.

The remainder of this chapter is divided into three main sections. In the first, we briefly discuss three theoretical models for understanding the role of strategic processing in learning, comprehension, and problem solving, with an eye to how strategic processing has been measured within these models. In the second, we describe, explain, illustrate, and problematize four different ways to assess strategic processing by means of self-reports. In the third, we summarize the results of our analysis and discuss implications and future directions.

Theoretical Background

We present and discuss three prominent models focusing on strategic processing. These models were developed within educational psychology decades ago but still influence thinking about strategies and research on strategic processing.

The Good Strategy User Model

This model describes three categories of strategies: goal- or task-specific strategies, monitoring strategies used to control and regulate goal-specific strategies, and higher-order strategies used to plan sequences of goal-specific and monitoring strategies (Pressley, 1987; Pressley, Borkowski, & Schneider, 1987). While the monitoring and planning strategies are considered forms of metacognitive procedures, metacognitive knowledge is represented in the model as specific strategy knowledge, including knowledge about how, when, and where to use particular strategies. Such knowledge also has a motivational aspect because it includes understanding that success is due to the use of task-appropriate strategies and that failure might have been avoided by the use of such strategies. In addition, general metacognitive knowledge about strategies has motivational properties because it includes understanding that strategic effort generally increases the likelihood of success. Finally, the good strategy user has an extensive knowledge base (in addition to the knowledge of and about strategies) that sometimes may make strategic effort superfluous and sometimes prompt strategy use but, most importantly, enables use of particular strategies (Pressley, 1987; Pressley et al., 1987).

Perhaps the most controversial aspect of the Good Strategy User Model is that it posits that good strategy use is often characterized by automaticity, implying that the components described above, including strategic processing, have been automatized or habituated and therefore do not require conscious attention. As noted above, such automatic processing tends to be categorized as skills rather than strategies in more contemporary theorizing.

Because the research leading up to the Good Strategy User Model overwhelmingly relied on randomized experiments, in which participants were trained to perform particular strategies and the effects on outcome variables such as recall of information were measured, the measurement of strategies per se was not a major issue. However, as noted by Levin (2008), Pressley was also eager to use qualitative methods, such as interviews or cued recall techniques, to gain insights into participants' thinking about their cognitive processing (e.g., Pressley & Levin, 1977), and he later became a strong proponent of verbal protocol analysis (e.g., Pressley & Hilden, 2004).

The Model of Domain Learning

This model describes the interplay of knowledge, strategies, and interest across three stages of academic development within a domain, which are termed acclimation, competence, and proficiency/expertise (Alexander, 1997, 2004, 2005, 2012). In addition to changes in the configuration of these main components that occur across the stages, the model acknowledges that there are many phases or episodes of learning within each stage, with these phases or episodes also characterized by certain interplays among knowledge, strategic processing, and interest (Alexander, 1997).

Regarding strategic processing, strategies at different levels of specificity (i.e., domain-general vs. domain-specific) and processing (i.e., surface- vs. deep-processing) are distinguished, as well as cognitive and metacognitive strategies (Alexander, 2004, 2005). Importantly, strategies are conceived of as effortful, intentional, and purposeful procedures directed toward improving learning, comprehension, and problem solving through the acquisition, transformation, and transfer of information (Alexander, 1997, 2004). While learners' dependency on surface strategies to establish a rudimentary knowledge base, gain foundational understanding, and solve elementary problems decreases during the acclimation stage, essentially levels off during the competence stage, and then again decreases during the proficiency/expertise stage, their use of deeper strategies to organize, transform, and critically analyze information increases across the three stages of domain learning and becomes particularly important in the service of knowledge generation, deep comprehension, and problem formulation during the proficiency/expertise stage (Alexander, 2004, 2012).

Of note is that combined shifts in knowledge and interest across the stages are seen as important contributors to the development and increased use of deeper processing strategies (Alexander, 1997). In effect, mutually influential relationships among a large, well-integrated body of knowledge, high individual interest, and a well-established and efficient repertoire of deeper processing strategies are regarded as a hallmark of the most advanced learners in a domain (Alexander, 1997, 2004).

With respect to the measurement of strategies, Dinsmore, Hattan, and List (2018) documented that the research conducted within the Model of Domain Learning overwhelmingly has measured strategic processing by means of offline self-report inventories. However, these offline self-reports have typically been task-specific and collected immediately after task completion (Alexander & Murphy, 1998; Alexander, Murphy, Wood, Duhon, & Parker, 1997; Alexander, Sperl, Buehl, & Fives, 2004). For example, Alexander et al. (1997) asked participants to monitor the strategies they used during task completion (i.e., text reading) and immediately afterwards check any strategy they had used to comprehend and remember the text on a list of 20 text-processing strategies, also marking the strategies they had found most helpful. As Dinsmore et al. (2018) showed, the reliability estimates reported for the kind of strategy measures used within the Model of Domain Learning have quite often been lower than desirable, and the relationship between scores on such measures and performance actually seems open to question.

The Cyclical Model of Self-Regulated Learning

This model describes the three cyclical phases of forethought, performance, and self-reflection (Zimmerman (2000, 2013). Forethought includes task analysis and self-motivation used in preparation for efforts to learn. The performance phase involves the execution and monitoring of strategies planned during the forethought phase. Finally, during self-reflection, individuals self-evaluate their learning and reflect on the causes of the outcome, as well as react emotionally to what happened during performance and draw inferences regarding future learning. The cyclical nature of the model involves that processes in one phase influence processes in the next, and that processes during self-reflection influence processes in the forethought phase when individuals continue their efforts to learn (Zimmerman, 2000, 2013).

Within Zimmerman’s model, self-regulated strategies are conceived of as purposefully selected or planful cognitive processes and behavioral actions directed at acquiring or displaying knowledge and skills. For example, strategies can facilitate learning and performance by helping students attend to, analyze, and reorganize academic tasks (Zimmerman, 2000). Of note is also that strategies are considered context-specific within self-regulated learning theory, implying that self-regulated students adjust their strategic choices and activities to different study contexts (Zimmerman, 2000). Finally, Zimmerman's view on self-regulated learning strongly emphasizes that strategic competence is of little value if individuals cannot motivate themselves to use it, with one key source of motivation being their self-efficacy perceptions.

Zimmerman's research on the identification and measurement of self-regulated strategies is strongly associated with the Self-Regulated Learning Interview Schedule (Zimmerman & Martinez-Pons, 1986, 1988, 1990). This methodology consists of a 15-min individual structured interview during which students are presented with different hypothetical learning contexts (e.g., when completing writing assignments outside class). For each context, students are asked to describe the methods they would use, and if they mention one or more strategies for a learning context, they are also asked to rate the frequency with which each mentioned strategy is used. Zimmerman and Martinez-Pons (1986, 1988, 1990) have shown substantial positive correlations between students' reports of strategies on this schedule and their academic achievement. However, in addition to using such offline self-reports about hypothetical learning contexts, Zimmerman and colleagues (Cleary & Zimmerman, 2001; Kitsantas & Zimmerman, 2002) have more directly evaluated processes included in the cyclical model by asking individuals close- and open-ended questions during actual task completion. We will return to this methodology in the section termed Task-Specific Self-Report Inventories.

Summary

To summarize, the three reviewed models provide foundational understanding of the importance and functioning of strategic processing. Taken together, they describe how strategic processing, functioning within a system of cognitive, metacognitive, and motivational components, can improve learning, comprehension, or problem solving. The models do, however, differ with respect to their adoption of a developmental perspective and the way they conceptualize strategic processing. Thus, while Pressley (1987; Pressley et al., 1987) certainly did not disregard the importance of development, the Good Strategy User Model is not a developmental model of strategic processing. By comparison, both the Model of Domain Learning and the Cyclical Model of Self-Regulated Learning could be described as developmental. These two models differ in terms of the grain sizes on which they focus, however, with the former focusing on comprehensive stages of academic development and the latter focusing on more fine-grained phases that unfold sequentially in a recursive manner over continued efforts to learn. With respect to the conceptualization of strategic processing, Pressley's model differs from the others because it allows for automatic strategic processing, and Alexander's model is unique in distinguishing between surface- and deep-processing strategies. Finally, we note that research conducted within the three models has largely relied on self-reports of strategic processing. In the following section, we further elaborate upon some of these self-report methodologies, ranging from thinking aloud during task completion to answering questions about task-specific processing retrospectively.

Measuring Strategic Processing through Concurrent and Task-Specific Self-Reports

Self-reports of strategic processing can involve on-line or off-line reporting. On-line approaches refer to measurements taken concurrent to task performance, such as having learners think aloud to create a verbal protocol that subsequently can be analyzed by the researchers. Off-line approaches refer to self-report inventories or interviews administered before or after task performance. In this chapter, we consider on-line measures in the form of concurrent thinking aloud and off-line measures in the form of task-specific self-reports collected immediately after task performance.

Verbal Protocol Analysis

In research on text-based learning and comprehension, individuals can be instructed to think aloud as they read one or more texts (Ericsson & Simon, 1993; Pressley & Afflerbach, 1995;

Trabasso & Magliano, 1996). This means that readers are asked to report whatever thoughts come to mind as they read, and the intent is to have them report only those thoughts that are immediately accessible and reportable in language, and therefore represent the contents of working memory (Ericsson & Simon, 1993). This activity can be considered to reflect an effortful search for meaning and, as such, to provide a window on strategic processing (Trabasso & Magliano, 1996). More specifically, verbal protocols resulting from concurrent thinking aloud are considered a valid measure of the metacognitive states that arise during reading and the strategies that respond to these states (Pressley & Afflerbach, 1995), in particular inference processes that support the construction of a mental model (Trabasso & Magliano, 1996). In this section, we describe two approaches to collecting verbal protocol data and discuss available evidence that these approaches are valid measures of comprehension strategies.

Approaches to collecting verbal protocols. One approach is to allow individuals to self-select when and where they think aloud while reading (Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Pressley & Afflerbach, 1995). This approach is viewed as sensitive to metacognitive states that arise during reading and strategies readers employ in response to them (e.g., experiencing confusion after reading a section and then rereading that section in response to that metacognitive state; Pressley & Afflerbach, 1995). That is, readers likely report thoughts that reflect an effortful search for meaning, especially when experiencing challenges understanding the texts.

An alternative approach is to have readers report their thoughts after pre-selected locations, which could be after every sentence, paragraph, or section (Trabasso & Magliano, 1996), or after theoretically determined locations (Kaakinen & Hyönä, 2005; Magliano & Millis, 2003). This approach can be construed as a retrospective protocol because participants are asked to reflect upon cognitive states that have recently occurred but likely remain in working memory (Ericsson & Simon, 1993). Typically, the intent is to reveal thoughts that pertain to mental model construction, in particular (Magliano, 1999; Trabasso & Magliano, 1996). Thus, this approach aims to reveal the products of knowledge activation and inference generation to integrate text constituents in the mental model. This may involve establishing how text constituents are semantically connected (e.g., cause and effect, claim-evidence, and contrastive relationships; Magliano, Trabasso, Graesser, 1999; Ray & Magliano, 2015) or integrating text constituents with background knowledge (Todaro, Magliano, Millis, McNamara, & Kurby, 2008). This approach is suitable for gaining access to processes involved in mental model construction for at least two reasons. First, the sentence prior to a think aloud prompt likely serves as a retrieval cue for knowledge that support mental model construction, which constrains the contents of working memory that are reported at the prompt. Presumably, these contents reflect the products of the processes that support mental model construction. Second, the instructions typically used with this approach may invoke a strategy to focus on thoughts that are reflective of mental model construction.

Approaches to coding verbal protocols. Verbal protocols can be used to answer research questions regarding strategic processing to the extent that a coding system is developed that is aligned with those research questions. A coding system can be developed through an inductive process that arises from qualitative methodologies, or it can be developed a priori, grounded in theory. In this section, we discuss a few notable examples of these different approaches.

Pressley and Afflerbach (1995) reviewed an extensive list of studies that employed an open-ended prompt methodology (i.e., readers chose when to think aloud) for collecting verbal protocols and a traditional, qualitative analysis. These studies suggested that skilled readers regularly monitor cognitive states as they read, and engage in a number of strategies in response to challenges they face in terms of learning from and comprehending texts (e.g., rereading, paraphrasing, elaborative inferencing).

Trabasso and Magliano (1996) provided an example of developing a coding system a priori inspired by theory. They had college students think aloud after each sentence of short narratives. The coding system was based on the constructionist theory of comprehension (Graesser, Singer, & Trabasso, 1994) and distinguished between three broad categories of inferences: explanations, predictions, and associations. Consistent with the theory, Trabasso and Magliano (1996) found that explanations predominated when examining the inferences produced while thinking aloud. Moreover, explanations based on prior text information provided the primary basis for establishing how distal sentences were related, supporting global coherence. The results were thus consistent with the assumption that comprehension is achieved largely through explanatory reasoning (e.g., Graesser et al., 1994).

A number of researchers have developed coding systems sensitive to such theoretically based distinctions, including a broader range of strategies that can be reflected in verbal protocols (Kaakinen & Hyönä, 2005; Magliano & Millis, 2003; Magliano, Millis, the RSAT Development Team, Levinstein, & Boonthum, 2011; McCarthy & Goldman, 2019; McMaster et al., 2012; Rapp, van den Broek, McMaster, Kendeou, & Epsin, 2007; Todaro et al., 2008). These strategies concern paraphrases, episodic recollections, bridging inferences, elaborative inferences, metacognitive judgments, evaluative statements, and affective judgments. Theoretically derived strategies that directly support the construction of a mental model (e.g., bridging inferences, elaborations) tend to be more frequent in verbal protocols than are other strategies (Rapp et al., 2007; Todaro et al., 2008).

However, the frequencies of different strategies vary with reading proficiency (Magliano & Millis, 2003; McMaster et al., 2012). For example, Magliano and Millis (2003) had college students read simple narrative texts and think aloud at selected sentences that afforded bridging inferences based on an a priori, theoretical analysis of the causal structure underlying the story events. They found that proficient readers tended to produce more bridging inferences than less proficient readers, whereas less proficient readers tended to produce more paraphrases than proficient readers. Rapp et al. (2007) and McMaster et al. (2012) compared different profiles of struggling middle school readers who thought aloud at every sentence while reading. While one profile group of struggling readers tended to paraphrase the sentence that was just read, another profile group primarily produced invalid elaborative inferences, both approaches indicating difficulties using deep-processing strategies.

Validating verbal protocols. Despite arguments that thinking aloud is minimally subject to task demands (Ericsson & Simon, 1993), it is possible that participants produce thoughts that would not occur during silent reading. Especially, when participants are asked to reconstruct their thoughts, the methodology is prone to fabrication. As such, it is crucial to adopt approaches that help identify which aspects of verbal protocols likely support comprehension and text-based learning in contexts where readers are not asked to think aloud. There are at least three approaches to validating verbal protocols.

The first approach is to demonstrate that the strategies resulting from verbal protocol analysis are correlated with measures of individual differences known to affect comprehension, such as proficiency in reading (McMaster et al., 2012) and comprehension (Kopatich, Magliano, Millis, Parker, & Ray, in press; Magliano & Millis, 2003; Magliano et al., 2011; Millis, Magliano, & Todardo, 2006), working memory capacity (Whitney, Ritchie, & Clark, 1991), proficiency in a second language (Zwaan & Brown, 1996), or disciplinary expertise (Graves & Frederiksen, 1991). The studies of Magliano and Millis (2003) and McMaster et al. (2012) cited above exemplify this type of validation. Zwaan and Brown (1996) provided a notable example with respect to demonstrating differences in strategic processing as a function of proficiency in a second language. They recruited participants that varied in knowledge of French and had them think aloud (in their native English) while reading narratives written in French or English. While reading in English, both groups of students demonstrated results consistent with Trabasso and Magliano (1996), with the most dominant strategy being explanation. While reading in French, however, only the proficient French readers demonstrated an explanation-based strategy, whereas the less proficient readers engaged in strategies that involved identifying the meaning of lexical items and syntactic relationships. As a final example of this approach, low working memory span readers have been shown to engage in excessive elaboration when thinking aloud compared to high span readers (Whitney et al., 1991).

A second approach to validating verbal protocol analysis is to demonstrate that strategies revealed by thinking aloud are correlated with measures of comprehension outcomes (Goldman et al., 2012; Kopatich et al., in press; Magliano & Millis, 2003; Magliano et al., 1999, 2011). For example, Magliano, Millis, and colleagues have shown that measures of bridging inferences are positively correlated with comprehension outcomes for texts used in collecting the verbal protocols (Magliano & Millis, 2003; Magliano et al., 2011; Millis et al., 2006), as well as for other texts (Magliano & Millis, 2003; Magliano et al., 2011). Conversely, the measures of paraphrasing tend to be negative correlated with comprehension performance (Magliano & Millis, 2003). Magliano and Millis (2003) argued that more proficient readers engage in strategies that promote coherence building, whereas less proficient readers engage in locally focused strategies directed at understanding individual sentences. Kopatich et al. (in press) found that measures of bridging and elaborative inferences partially mediated the effects of general and language processing resources on comprehension outcomes.

A final approach for validating verbal protocols is to demonstrate that they are correlated with moment-to-moment measures of reading, which would suggest that these processes occur during silent reading. The study by Magliano et al. (1999) is an example of this approach. In that study, the explanations, predictions, and associations included in the verbal protocols of one group of students predicted variations in reading times in another group of students who read the same texts silently under different goal conditions (e.g., reading to explain vs. reading to predict). As another example, Kaakinen and Hyönä (2005) collected verbal protocols and eye movement data while participants read expository texts, examining the length of fixations at sentences where participants were asked to think aloud. Those authors found significant correlations between the verbal protocols and eye movement data, with questions and explanations in the protocols, in particular, being positively correlated with fixation times.

Eye Movement Cued Self-Reports

One novel approach to eliciting self-reports of strategic processing is to record participants' eye movements while they are peforming a task, and afterwards display those eye movements to them to cue reports of their thinking during task performance. Eye movement cued self-reports can be considered a form of retrospective thinking aloud because participants are responding to external cues rather than internal cognitive states. Thus, they are retrospective in the sense that participants are asked to reflect upon strategic processing in the near past, with the products of those strategies likely still in working memory. To elicit these reports, participants are presented with a video of what they just saw while performing a specific task, with the video superimposing participants' eye movements as a small circle indicating where they were looking during task performance. As eye movements are considered to reflect cognitive processes (Rayner, 2009), the rationale is that they may cue participants to report orally on such processes as they were taking place in a specific task environment. In particular, eye movements are assumed to represent an externalization of strategic processes that may be used to cue descriptions of those processes.

Eye movement cued self-reports have been used to measure strategic processing in a variety of learning tasks, including multimedia learning (de Koning, Tabbers, Rikers, & Paas, 2010; Stark, Brünken, & Park, 2018), text reading (Catrysse, Gijbels, & Donche, 2018; Penttinen, Anto, & Mikkilä-Erdmann, 2013; Salmerón, Naumann, García, & Fajardo, 2017), web search (Brand-Gruwel, Kammerer, van Meeuwen, & van Gog, 2017; Muntinga, & Taylor, 2018), visual tasks (Jarodzka, Scheiter, Gerjets, & van Gog, 2010), mathematical problem solving (Rau, Aleven, & Rummel, 2017; van der Weijden, Kamphorst, Willemsen, Kroesbergen, & van Hoogmoed, 2018), and interaction with and inspection of complex systems (Ruckpaul, Fürstenhöfer, & Matthiesen, 2015; van Meeuwen, Brand-Gruwel, Kirschner, de Bock, & van Merriënboer, 2018; van Gog, Paas, van Merriënboer, & Witte, 2005).

Issues concerning the presentation of eye movements. A critical aspect of eye movement videos concerns speed. As eye movements tend to be quick, participants easily may lose track of the fixations. Therefore, Russo (1979) recommended that the replay proceed at a rate compatible with recall and reports. Accordingly, several researchers have used slower speed for video replays, such as 50 (Brand-Gruwel et al., 2017; de Koning et al., 2010; Hyrskykari, Ovaska, Majaranta, Räihä, & Lehtinen, 2008) or 75% of the actual speed (van Meeuwen et al., 2018). Others have replayed the videos at full speed, however (Jarodzka et al., 2010; Salmerón et al., 2017; Stark et al., 2018). While it is evident that eye-movements are faster than verbalizations, it is not so obvious that each fixation represents a unique processing episode that deserves to be commented on. Accordingly, Guan, Lee, Cuddihy, and Ramey (2006) reported that 47% of eye movement visits to particular regions of interest did not elicit any verbal comments. This is especially clear when performing reading tasks, where participants may engage for minutes in reading a long paragraph. Slowing down the replay of this kind of episode may not provide new insights into participants' ongoing processing because they are tapping the same process; rather, it may induce negative reactions toward the task. Researchers also have varied the way videos are replayed as a way to accommodate this issue. Thus, participants have been allowed to change the speed of the recordings (Russo, 1979) or to pause the videos at any time (de Koning et al., 2010; van Gog et al., 2005). Such procedures allow students to report complex thoughts in tasks where a sequence of fixations reflects different cognitive processes. In other cases, such as when a long sequence of fixations reflects a single process (e.g., deep reading), the researcher can just pause the video at points of critical interest (Penttinen et al., 2013; Ruckpaul et al., 2015).

Another relevant issue concerns how eye-movements are represented in the videos. While the major components of eye-movements include both fixations and saccadic movements, the majority of studies have presented eye-movements only as individual fixations, with the size of the circles representing the length of the fixations. During task performance, previous fixations disappear and new ones appear in different locations. This way of representing eye movements thus emphasizes location (i.e., fixation) over orientation (i.e., saccadic movement). It is therefore possible that it increases the likelihood that processing related to objects (e.g., describing objects) are elicited, compared to more dynamic processing (e.g., comparing and contrasting objects). In a rare exception, Ruckpaul et al. (2015) replayed videos representing eye movements during task performance as fixations (i.e., circles) as well as saccadic movements (i.e., lines between circles). Unfortunately, these authors did not include a condition showing only fixations, which made it impossible to address the possibility mentioned above.

Validating eye movement cued self-reports. Regarding validity, eye movement cued self-reports have been proposed as a less intrusive alternative to concurrent thinking aloud (van Gog et al., 2005). However, as noted earlier, given that participants are asked to reconstruct their thoughts, the methodology is prone to fabrications. Such fabrication may still be less frequent than for retrospective self-reports that do not use eye-movement cues because the reconstruction, at least, must be consistent with the sequence of fixations (Guan et al., 2006).

In the fields of usability and learning research, there have been some attempts to analyze the reliability and validity of this methodology by comparing it to concurrent thinking aloud. Not surprisingly, thinking aloud in response to cues provided by eye movements tends to involve a higher number of verbalizations (Brand-Gruwel et al., 2017; Hansen, 1991; Hyrskykari et al., 2008; Russo, 1979). This may be due, in part, to the fact that participants spend more time reporting when cued by eye movements. Once reporting time is controlled for, however, this difference seems to disappear (Ruckpaul et al., 2015). More importantly, the two methodologies may elicit different types of thoughts (Brand-Gruwel et al., 2017; Hansen, 1991; Hyrskykari et al., 2008; Ruckpaul et al., 2015; van Gog et al., 2005). In usability studies, where participants are assessed as they interact with a system (e.g., when buying products from different web pages), the percentage of manipulative comments (e.g., "I write the name into this field") is lower with eye movement cued than with concurrent thinking aloud, while the difference in terms of cognitive comments (e.g., reflections about the system) is less clear (Hansen, 1991; Hyrskykari et al., 2008). In studies of learning, where participants are engaged in a task to meet a learning goal, the major difference between the two methodologies seems to be the number of metacognitive and evaluative comments produced (i.e., evaluations of the information and the learning process), with higher frequencies observed in eye movement cued verbal protocols (Brand-Gruwel et al., 2017; van Gog et al., 2005). In brief, such comparative studies suggest that protocols from concurrent thinking aloud reflect more processing concerning what participants are doing, while protocols from eye movement cued thinking aloud reflect more processing involving evaluation of their actions.

As mentioned earlier, a major threat for eye movement cued thinking aloud is fabrication. Guan et al. (2006) compared the sequences of verbalizations in such protocols to the actual sequences of problem solving as indicated by participants' eye movements. More than 80% of participants' verbalizations of what they were attending to corresponded to the actual sequences of their eye movements, and less than 3% of the verbalizations corresponded to areas of the task scenario that were not identified by the eye movement data. This pattern of results suggests a low rate of fabrication, but this point should be taken with caution. To the best of our knowledge, no other study has empirically tested the issue of fabrication with eye movement cued self-report data.

Thus far, there also have been few attempts to validate this methodology by testing its predictive power. Elling, Lentz, and de Jong (2011) found that participants thinking aloud concurrently or retrospectively in response to eye movements did not differ in terms of the usability problems detected. Brand-Gruwel et al. (2017) reported a global positive correlation between evaluative verbalizations and performance (i.e., selection of appropriate web pages), but correlations were not reported separately for each methodology (i.e., concurrent and retrospective eye movement cued thinking aloud). Thus, any claim about the validity of this methodology should be considered with caution until further documentation is available.

Task-Specific Self-Report Inventories

Some self-report inventories measure strategic processing as an aptitude or a trait, by asking respondents "to generalize their actions across situations rather than referencing singular and specific learning events" (Winne & Perry, 2000, p. 542). In a review, Veenman (2005) showed that correlations between strategy data from such self-report inventories and on-line methodologies, such as concurrent thinking aloud, are generally low, indicating poor convergent validity due to several issues with off-line self-report inventories, such as their decontextualized and retrospective nature. However, as posited by Schellings (2011), low correlations between strategy data from self-report inventories and on-line methodologies may not necessarily indicate the invalidity of data gathered by means of self-report inventories, but reflect the lack of specificity and proximity of many self-report inventories to actual task contexts. Accordingly, she showed that self-reports of strategies on an inventory that was tailored to a particular reading task context yielded scores that were fairly well aligned with on-line strategy data.

Of note is that such task-specific strategy inventories ask learners to make judgments about their strategic activities in a specific task, rather than about what they generally or typically do during learning or comprehension. The logic of this approach is also consistent with Ericsson and Simon’s (1993) recommendation that retrospective verbal protocols be provided immediately after task performance, that is, when metacognitive states and strategic activities are still relatively accessible for retrieval in working memory.

In several studies, Bråten and colleagues (Anmarkrud & Bråten, 2009; Bråten & Anmarkrud, 2013; Bråten & Samuelstuen, 2004, 2007; Samuelstuen & Bråten, 2005, 2007), in the context of learning from a single expository text, have demonstrated the possibility to obtain valid measurements when using self-report strategy inventories with items tailored to a specific task context and administered shortly after task performance. These authors have followed the four guidelines for constructing task-specific strategy inventories explicated by Bråten and Samuelstuen (2007). First, a specific task (e.g., to read a text for a particular purpose) must be administered, to which the items on the inventory are referring. Second, the task must be accompanied by an instruction that directs learners to monitor their strategies during task performance and informs them that they will be asked some questions afterwards about how they proceeded. Third, to minimize the retention interval, the strategy inventory must be administered immediately after task completion. Finally, in referring to recent episodes of strategic processing, the wordings of task-specific items must be different from more general statements. This means that general item stems such as “when I study” or “when I read” must be omitted in task-specific items. Moreover, to make it clear that the items refer back to the recently completed task, the verb must be in the past tense (e.g., “I tried to understand the content better by relating it to something I know”).

Validating task-specific self-report inventories. Strategy scores on a task-specific self-report inventory based on these guidelines have been validated in several ways. For example, Bråten and Samuelstuen (2007) showed that strategies self-reported in this way corresponded quite closely with strategies traced in the study materials, with correlations between self-reported and traced surface-level strategies exceeding .75 and correlations between self-reported and traced deeper-level strategies exceeding .80. Moreover, Samuelstuen and Bråten (2005, 2007) demonstrated that students’ strategy scale scores accounted for their performance on expository text comprehension tasks, with correlations between deeper-level strategies and performance exceeding .35 (see also, Bråten & Anmarkrud, 2013).

The same type of task-specific self-report inventory was developed by Bråten and Strømsø (2011) to measure strategic processing when individuals read multiple texts on the same topic. Specifically, this measure was constructed to assess a surface strategy involving the accumulation of pieces of information from different texts and a deeper-level strategy involving cross-text elaboration. These two dimensions of multiple text comprehension strategies have been confirmed through factor analysis and scores on these dimensions have been found to predict performance (Bråten, Anmarkrud, Brandmo, & Strømsø, 2014; Bråten & Strømsø, 2011), with the accumulation strategy negatively and the cross-text elaboration strategy positively related to intertextual comprehension performance. Further validation of scores on this inventory was provided by Hagen, Braasch, and Bråten (2014), who compared strategy scale scores with strategies as revealed by students’ spontaneous note-taking, and by

List, Du, Wang, and Lee (2019), who compared strategy scale scores with multiple text model construction as revealed by students' written responses and with integrative processing as revealed by their think-aloud utterances.

Although valid scores have been obtained on task-specific self-report inventories of strategic processing administered immediately after task performance, such measures may be subject to some of the same errors that seem to plague general, decontextualized strategy inventories. Thus, because individuals in any case have to retrieve strategic episodes after some delay, fallible, biased, and reconstructive memory processes cannot be ruled out. In particular, the social desirability of response alternatives may bias people's self-reports of strategies. Accordingly, people may report more strategic processing than actually executed because they believe this is to be approved by others. It is also possible that people report much use of certain strategies because they believe that those strategies are effective, not because they actually used them to any great extent. Yet another potential problem with task-specific self-report inventories administered after task performance may be a tendency to report using strategies described by the items although other strategies were actually used.

Questioning during task performance. Given such potential threats to the validity of scores on task-specific inventories administered after task performance, the procedure used by Zimmerman and colleagues (Cleary & Zimmerman, 2001; Kitsantas & Zimmerman, 2002) may be a viable option. In this approach, participants have been asked a series of task-specific questions about their self-regulatory processes, including their strategies, not only before and after but also during their efforts to learn particular athletic skills. While some of these questions, for example regarding self-efficacy, were close-ended with participants responding on a scale from 0 to 100, others were open-ended with responses coded by the researchers. For example, Kitsantas and Zimmerman (2002), who studied the practice of volleyball overhand serving among college women, stopped the participants during the practice episode and asked them about their strategy for successfully executing the next overhand serve after having missed the target on the two preceding attempts. The self-reported strategies were then coded into the categories of specific techniques, visualization strategies, concentration strategies, both specific techniques and concentration strategies, and practice/no strategy.

This approach was found to be valid in the sense that it differentiated between participants' at different levels of expertise with respect to the processes included in the three phases of the cyclical model of self-regulated learning (Zimmerman, 2000, 2013). In the performance phase, in particular, the experts displayed more use of specific technique strategies than did other participants and also monitored their use of specific techniques and outcomes to a greater extent. Further, self-regulatory processes measured in this way were highly correlated with participants' serving performance after the practice (Kitsantas & Zimmerman, 2002). Similar results were obtained by Cleary and Zimmerman (2001), who used task-specific questioning to study self-regulatory processes during participants' practice on a particular basketball skill.

Diary Methods

Diary methods are systematic and structured ways of measuring psychological variables repeatedly in participants' natural learning environment (Iida, Shrout, Laurenceau, & Bolger, 2012). As measures of strategic processing, diaries can be considered standardized instruments used to self-report or self-record strategic activities on a daily basis near the time and in the contexts they occur (Schmitz, Klug, & Schmidt, 2011). While the most common diary method design involves responding to a series of questions about one's own activities once at the same time every day, such recordings sometimes take place several times a day and, at a minimum, after a few days (Iida et al., 2012; Schmitz et al., 2011).

Because diaries are completed in participants' daily environment, diary data can be said to fare well in terms of ecological validity compared to data collected in experimental settings. Moreover, compared to task-specific self-report inventories, which typically refer to one single task context at a single point in time, diaries allow for the measurement of strategies in a range of contexts over time. Because diary methods allow for the aggregation of participant responses over time, they may lead to more valid information than one-shot strategy measurements (Iida et al., 2012). That learners record their strategic activities on a daily basis with reference to specific contexts taking place the same day, may also reduce the chances of retrospective biases (Tourangeau, 2000) and, consequently, increase the validity of such self-reports.

Diary formats. According to Iida et al. (2012), there are three commonly used diary formats: paper-and-pencil diaries, brief telephone interviews, and electronic diaries. Electronic response formats have become the most common, with participants logging into a secure website daily to complete an online questionnaire, with researchers reminding participants by means of electronic devices if responses are not entered, and with questionnaire data easily transferred to computer programs for statistical analysis.

As an example of a standardized web-based diary in the area of strategic processing, Andreassen, Jensen, and Bråten (2017) constructed a diary that referred to three study contexts: attending lectures, individual study (i.e., studying alone), and social study (i.e., studying with others). Based on a review of the literature on self-regulated study strategies, they identified 20 strategies reportedly used in these three contexts, with six strategies referring to the context of attending lectures (e.g., asking questions to the lecturer during or after lectures), eight strategies referring to the context of individual study (e.g., making drawings or figures to better understand text), and six strategies referring to the context of social study (e.g., consulting fellow students). For each study context that a participant had participated in on a particular day, the participant also recorded in the diary whether he or she had used the strategies associated with that context. For each strategy that was recorded in each study context, the perceived benefit of that strategy was also recorded on a Likert-type scale. Before the data collection started, Andreassen et al. (2017) presented and demonstrated the web-based diaries to the participants, and also had them practice accessing and completing the diaries on their computers or smartphones. During the data collection, which lasted for 12 consecutive days, participants' daily recordings on the diary websites were tracked by the researchers, who contacted every participant who had not completed the diary in the evening with a reminder (via SMS) to enter the diary data for that day.

Validating diary data. Several studies have indicated that diary methods may yield valid data on learning processes. For example, Kanfer, Reinecker, and Schmelzer (1996) found high accuracy of diary data when correlating them with data from external observers, and Schmitz and Wiese (2006) demonstrated the effectiveness of an intervention to promote self-regulated learning by means of diary data collected over 35 days. In the Andreassen et al. (2017) study, differences in the reported use and benefits of particular visual and social strategies were observed between college students with and without dyslexia. Still, there are several threats to validity when applying diary methods.

First, depending on the length of the diary questionnaire, the frequency of diary completion, and the length of the diary period, diary methods may burden participants and lead to lack of motivation, non-compliance, and attrition (Iida et al., 2012). Presumably, diary methods are more vulnerable to lack of compliance than other self-report methods because data are essentially collected by the participants themselves, without any researchers present.

Second, completing diaries may produce reactivity effects that confound measurements (Schmitz et al., 2011). That is, daily self-monitoring and reporting may change participant activities, for example, lead to an increase in strategy use because students become more aware of and motivated to change their strategies in desired directions. Moreover, such reactivity may increase when recordings are temporally close to the recorded activities

(Schmitz et al., 2011). Thus, on the one hand, recordings that are temporally close to the activities may reduce retrospective biases; on the other hand, they may compromise validity because of heightened reactivity. According to Iida et al. (2012), however, repeated completion of diaries may lead to a form of habituation that over time will reduce reactivity.

Third, a potential threat to validity concerns construct variability. That is, during a diary period, participants' understanding of a particular construct may change. For example, a participant may start out with a narrow understanding of what constitutes "constructing a summary," but through repeated self-monitoring and reflection and exposure to a diary strategy questionnaire, he or she may come to understand this construct more broadly, with implications for how such strategic activities are recorded in the diary.

Finally, as with task-specific self-report inventories, self-reporting strategic processing on a diary questionnaire where certain strategies are prelisted likely will restrict strategy measurements to those strategies. Thus, although diaries also may allow participants to report on strategies that are not prelisted (Andreassen et al., 2017), they may be less likely to do so because the prelisting makes them more aware of those strategies and less sensitive to other forms of strategic processing that they actually may engage in.

Conclusions, Implications, and Future Directions

The theoretical and empirical work discussed in this chapter highlights the important role of strategic processing in learning, comprehension, and problem solving, as well as the need to measure this complex construct in a valid way. We reviewed three theoretical models that have strongly influenced thinking and research on strategic processing within educational psychology. This review clarified the pivotal role of strategic processing within an interactive system of cognitive, metacognitive, and motivational components. Moreover, it showed that the authors of these models, to a large extent, have relied on concurrent and task-specific self-reports in attempting to measure strategic processing. This is not peculiar, given the emphasis on strategies as effortful, conscious, and intentional activities in the service of learning, comprehension, and problem solving. Simply stated: If people are aware of what they do when striving to achieve a particular goal, why not ask them to report what they do, either while doing it or immediately afterwards? As our subsequent discussion of such self-report methodologies showed, however, there are lingering issues with some of these methodologies that may represent challenges for the interpretation of data and replication of results.

First and foremost, a general lack of studies that compare strategy data across different methodologies makes it difficult to say with any degree of certainty whether what is captured (or missed) with respect to strategic processing is unique to particular methodologies. Thus, although data obtained with the concurrent and task-specific self-report methodologies that we discussed in this chapter to some degree have been compared with data from more objective methodologies, such as reading time, eye movements, and traces, such information is sparse and essentially lacking for most of the methodologies. Accordingly, this is an issue that needs further clarification. To involve meaningful comparisons, however, it is important that future researchers in this area try to ensure that more objective measurements (e.g., reading times or eye movements) actually reflect effortful strategic processing, as data from such methodologies are not self-explanatory and also may reflect bottom-up automatic processing or skills (Salmerón, Gil, & Bråten, 2018).

In addition, the different concurrent and task-specific methodologies that we discussed have seldom been compared with one another. Therefore, it is currently not known, for example, to what extent eye movement cued self-reports and task-specific self-report inventories administered immediately after task performance capture the same aspects of strategic processing, or to what extent task-specific self-report inventories and diary methods yield comparable data with respect to strategy use. Consequently, further research including more than one of these measures is needed to probe their concurrent validity.

Finally, different approaches to collecting data with each of the concurrent and task-specific self-report methodologies may yield different findings regarding strategic processing. For example, whether concurrent thinking aloud is self-directed or researcher-directed (i.e., prompted to occur at pre-selected locations), whether eye movement cued self-reports are based on fixations alone or both fixations and saccadic movements, whether task-specific self-report inventories are administered during or immediately after task performance, and whether diaries are completed once or several times a day may provide somewhat different windows on strategic processing. More systematic knowledge about how strategy data may differ with variants of the same type of self-report methodology is therefore also needed.

Taken together, the issues just discussed makes it hard to interpret and compare findings from different studies because it is difficult, if not impossible, to determine how much variation is due to researchers' methodological choices or preferences. Given current concerns about failures to replicate findings in psychological and cognitive science (Lindsay, 2015), this problem is not trivial. That is, as it now stands, failure to replicate findings across studies of strategic processing and how it independently and interactively contributes to learning, comprehension, and problem solving may, at least in part, be due to different approaches to strategy measurement across studies. Likewise, varying results of instructional efforts to improve strategic functioning across studies may be due to strategy measures that are differentially sensitive to such efforts or capture different aspects of strategic processing. Hopefully, this chapter will provide an impetus for much further multiple method research including one or more of the self-report methodologies that we discussed in combination with other types of strategy measures to clarify the similarities and differences between those measures.

References

Afflerbach, P., Pearson, P.D., & Paris, S.G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*, 364-373.

Alexander, P.A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M.L. Maehr & P.R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 213-250). Greenwich, CT: JAI.

Alexander, P.A. (2004). A Model of Domain Learning: Reinterpreting expertise as a multidimensional, multistage process. In D.Y. Dai & R.J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 273-298). Mahwah, NJ: Erlbaum.

Alexander, P.A. (2005). The path to competence: A lifespan development perspective on reading. *Journal of Literacy Research, 37*, 413-436.

Alexander, P.A. (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist, 47*, 259-280.

Alexander, P.A., Graham, S., & Harris, K.R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*, 129-154.

Alexander, P.A., & Murphy, P.K. (1998). Profiling the differences in students' knowledge, interest, and strategic processing. *Journal of Educational Psychology, 90*, 435-447.

Alexander, P.A., Murphy, P.K., Woods, B.S., Duhon, K.E., & Parker, D. (1997). College instruction and concomitant changes in students' knowledge, interest, and strategy use: A study of domain learning. *Contemporary Educational Psychology, 22*, 125-146.

Alexander, P.A., Sperl, C.T., Buehl, M.M., Fives, H., & Chiu, S. (2004). Modeling domain learning: Profiles from the field of special education. *Journal of Educational Psychology, 96*, 545-557.

Andreassen, R., Jensen, M.S., & Bråten, I. (2017). Investigating self-regulated study strategies among postsecondary students with and without dyslexia: A diary method study. *Reading and Writing, 30*, 1891-1916.

Anmarkrud, Ø., & Bråten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences, 19*, 252-256.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L., & van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of Computer Assisted Learning*, *33*, 234-251.

Bråten, I., & Anmarkrud, Ø. (2013). Does naturally occurring comprehension strategies instruction make a difference when students read expository text? *Journal of Research in Reading, 36*, 42-57.

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H.I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction, 30*, 9-24.

Bråten, I., & Samuelstuen, M.S. (2004). Does the influence of reading purpose on reports of strategic text processing depend on students' topic knowledge? *Journal of Educational Psychology, 96*, 324-336.

Bråten, I., & Samuelstuen, M.S. (2007). Measuring strategic processing: Comparing task-specific self-reports with traces. *Metacognition and Learning, 2*, 1-20.

Bråten, I., & Strømsø, H.I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning, 6*, 111-130.

Brown, A.L., Bransford, J.D., Ferrara, R.A., & Campione, J.C. (1983). Learning, remembering, and understanding. In J.H. Flavell & E.M. Markman (Eds.), *Handbook of child psychology. Volume III: Cognitive development* (pp. 77-166). New York: Wiley.

Catrysse, L., Gijbels, D., & Donche, V. (2018). It is not only about the depth of processing: What if *eye* am not interested in the text? *Learning and Instruction, 58,* 284-294.

Cleary, T., & Zimmerman, B.J. (2001). Self-regulation differences during athletic practice by experts, non-experts, and novices. *Journal of Applied Sport Psychology, 13,* 185-206.

de Koning, B.B., Tabbers, H.K., Rikers, R.M.J.P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20,* 111-122.

Dinsmore, D.L., Hattan, C., & List, A. (2018). A meta-analysis of strategy use and performance in the Model of Domain Learning. In H. Fives & D.L. Dinsmore (Eds.), *The Model of Domain Learning: Understanding the development of expertise* (pp. 37-55). New York: Routledge.

Elling, S., Lentz, L.R., & de Jong, M. (2011). Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1161-1170). New York: Association for Computing Machinery.

Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

Goldman, S.R., Braasch, J.L.G., Wiley, J., Graesser, A.C., & Brodowinska, K.M. (2012). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly, 47,* 356-381.

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.

Graves, B., & Frederiksen, C.H. (1991). Literary expertise in the description of a fictional narrative. *Poetics, 20*, 1-26.

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1253-1262). New York: Association for Computing Machinery.

Hagen, Å.M., Braasch, J.L.G., & Bråten, I. (2014). Relationships between spontaneous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. *Journal of Research in Reading, 37*, 141-157.

Hansen, J.P. (1991). The use of eye mark recordings to support verbal retrospection in software testing. *Acta Psychologica*, *76*, 31-49.

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.J., & Lehtinen, M. (2008). Gaze path stimulation in retrospective think-aloud. *Journal of Eye Movement Research, 2*, 1-18.

Iida, M., Shrout, P.E., Laurenceau, J.-P., & Bolger, N. (2012). Using diary methods in psychological research. In H. Cooper (Ed.), *APA handbook of research methods in psychology* (Vol.1, pp. 277-305). Washington, DC: American Psychological Association.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, *20*, 146-154.

Kaakinen, J.K., & Hyönä, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes, 40*, 239–257.

Kanfer, R., Reinecker, H., & Schmelzer, D. (1996). *Selbstmanagement-Therapie* [*Self-management therapy*] (2nd ed.). Berlin: Springer.

Kendeou, P., & O'Brien, E.J. (2018). Reading comprehension theories: A view from the top down. In M.F. Schober, D.N. Rapp, & M.A. Britt (Eds.), *The Routledge handbook of discourse processes* (2nd ed., pp. 7-21). New York: Routledge.

Kitsantas, A., & Zimmerman, B.J. (2002). Comparing self-regulatory processes among novice, non-expert, and expert volleyball players: A microanalytic study. *Journal of Applied Sport Psychology, 14*, 91-105.

Kopatich, R.D., Magliano, J.P., Millis, K.K., Parker, C.P., & Ray, M. (in press). Understanding how language-specific and domain-general resources support comprehension. *Discourse Processes*.

Levin, J.R. (2008). The unmistakable professional promise of a young educational psychology researcher and scholar. *Educational Psychologist*, *43*, 70-55.

Lindsay, D.S. (2015). Replication in psychological science. *Psychological Science, 26*, 1827-1832.

List, A., Du, H., Wang, Y., & Lee, H.Y. (2019). Toward a typology of integration: Examining the documents model framework. *Contemporary Educational Psychology*, *58*, 228-242.

Magliano, J.P. (1999). Revealing inference processes during text comprehension. In S.R. Goldman, A.C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 55-75). Mahwah, NJ: ErIbaum.

Magliano, J.P., & Millis, K.K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*, 251-283.

Magliano, J.P., Millis, K.K., The RSAT Development Team, Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition and Learning*, *6*, 131-154.

Magliano, J.P., Trabasso, T., & Graesser, A.C. (1999). Strategic processes during comprehension. *Journal of Educational Psychology*, *91*, 615-629.

McCarthy, K.S. & Goldman, S.R. (2019). Constructing interpretive inferences about literary text: The role of domain-specific knowledge. *Learning and Instruction, 60*, 245-251.

McMaster, K.L., van den Broek, P., Espin, C.A., White, M.J., Rapp, D.N., Kendeou, P., ... Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, *22*, 100-111.

Millis, K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and latent semantic analysis. *Scientific Studies of Reading*, *10*, 225-240.

Muntinga, T., & Taylor, G. (2018). Information-seeking strategies in medicine queries: A clinical eye-tracking study with gaze-cued retrospective think-aloud protocol. *International Journal of Human–Computer Interaction, 34*, 506-518.

Paris, S.G., Newman, R.S., & Jacobs, J.E. (1985). Social contexts and functions of children's remembering. In M. Pressley & C.J. Brainerd (Eds.), *Cognitive learning and memory in children* (pp. 81-115). New York: Springer.

Penttinen, M., Anto, E., & Mikkilä-Erdmann, M. (2013). Conceptual change, text comprehension and eye movements during reading. *Research in Science Education*, *43*, 1407-1434.

Pressley, M. (1986). The relevance of the Good Strategy User Model to the teaching of mathematics. *Educational Psychologist*, *21*, 139-161.

Pressley, M., & Afflerbach., P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.

Pressley, M., Borkowski, J.G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta (Ed.), *Annals of child development* (Vol. 4; pp. 89-129). Greenwich, CT: JAI Press.

Pressley, M., Forrest-Pressley, D.L., Elliott-Faust, D., & Miller, G. (1985). Children's use of cognitive strategies, how to teach strategies, and what to do if they can't be taught. In M. Pressley & C.J. Brainerd (Eds.), *Cognitive learning and memory in children* (pp. 1-47). New York: Springer.

Pressley, M., & Hilden, K. (2004). Verbal protocols of reading. In N.K. Duke & M.H. Mallette (Eds.), *Literacy research methodologies* (pp. 308-321). New York: Guilford.

Pressley, M., & Levin, J.R. (1977). Developmental differences in subjects' associative-learning strategies and performance: Assessing a hypothesis. *Journal of Experimental Child Psychology, 24*, 431-439.

Rapp, D.N., van den Broek, P., McMaster, K.L., Kendeou, P., & Espin, C.A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11*, 289–312.

Rau, M.A., Aleven, V., & Rummel, N. (2017). Making connections among multiple graphical representations of fractions: Sense-making competencies enhance perceptual fluency, but not vice versa. *Instructional Science, 45*, 331-357.

Ray, M**.,** & Magliano, J.P. (2015, November). *Automatic assessment of moment-to-moment comprehension processes*. Paper presented at the annual meeting for the Society for Computers in Psychology, Chicago, IL.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457-1506.

Ruckpaul, A., Fürstenhöfer, T., & Matthiesen, S. (2015). Combination of eye tracking and think-aloud methods in engineering design research. In J.S. Gero & S. Hanna (Eds.), *Design computing and cognition'14* (pp. 81-97). Cham, Switzerland: Springer.

Russo, J.E. (1979). A software system for the collection of retrospective protocols prompted by eye fixations. *Behavior Research Methods & Instrumentation*, *11*, 177-179.

Salmerón, L., Gil, L., & Bråten, I. (2018). Using eye-tracking to assess sourcing during multiple document reading: A critical analysis. *Frontline Learning Research, 6*, 105-122.

Salmerón, L., Naumann, J., García, V., & Fajardo, I. (2017). Scanning and deep processing of information in hypertext: An eye tracking and cued retrospective think-aloud study. *Journal of Computer Assisted Learning*, *33*, 222-233.

Samuelstuen, M.S., & Bråten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology, 46*, 107-117.

Samuelstuen, M.S., & Bråten, I. (2007). Examining the validity of self-reports on scales measuring students' strategic processing. *British Journal of Educational Psychology, 77*, 351-378.

Schmitz, B., Klug, J., & Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B.J. Zimmerman & D.H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 251-266). New York: Routledge.

Schmitz, B., & Wiese, B.S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31*, 64-96.

Shellings, G. (2011). Applying learning strategy questionnaires: Problems and possibilities. *Metacognition and Learning*, *6*, 91-109.

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A mixed-methods study using eye tracking. *Computers & Education, 120,* 185-196.

Todaro, S.A., Magliano, J.P., Millis, K.K., McNamara, D.S., & Kurby, C.A. (2008). Assessing the structure of verbal protocols. In B.C. Love, K.McRae, & V.M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 607-612). Austin, TX: Cognitive Science Society.

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. In A.A. Stone, J.S. Turkkan, C.A. Bachrach, J.B. Jobe, H.S. Kurtzman, & V.S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp. 29-47). Mahwah, NJ: Erlbaum.

Trabasso, T., & Magliano, J.P. (1996). Conscious understanding during text comprehension. *Discourse Processes, 21*, 255-288.

van der Weijden, F.A., Kamphorst, E., Willemsen, R.H., Kroesbergen, E.H., & van Hoogmoed, A.H. (2018). Strategy use on bounded and unbounded number lines in typically developing adults and adults with dyscalculia: An eye-tracking study. *Journal of Numerical Cognition, 4*, 337-359.

van Gog, T., Paas, F., van Merriënboer, J.J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*, 237, 244.

van Meeuwen, L.W., Brand-Gruwel, S., Kirschner, P.A., de Bock, J.J.P.R., & van Merriënboer, J.J.G. (2018). Fostering self-regulation in training complex cognitive tasks. *Educational Technology Research and Development, 66*, 53-73.

Veenman, M.V.J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Arelt & B. Moschner (Eds.), *Lernstrategien und*

Metakognition: Implikationen für Forschung und Praxis [*Learning strategies and metacognition: Implications for research and practice*] (pp. 75-97). Berlin: Waxman.

Veenman, M.V.J. (2011). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition and Learning*, *6*, 205–211.

Veenman, M.V.J. (2016). Metacognition. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 26-40). New York: Routledge.

Weinstein, C.E., Husman, J., & Dierking, D.R. (2000). Self-regulation interventions with a focus on learning strategies. In M. Boekaerts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 727-747). San Diego, CA: Academic Press.

Whitney, P., Ritchie, B.G., & Clark, M.B. (1991). Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*, *14*, 133-145.

Winne, P.H., & Perry, N.E. (2000). Measuring self-regulated learning. In M. Boekaerts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 532-568). San Diego, CA: Academic Press.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press.

Zimmerman, B.J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*, 135-147.

Zimmerman, B.J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *23*, 614-628.

Zimmerman, B.J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*, 284-290.

Zimmerman, B.J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Zwaan, R.A., & Brown, C.M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, *21*, 289-327.