How often are thoughts metacognitive? Findings from research on self-regulated learning, think-aloud protocols, and mind-wandering

By: Megan L. Jordano and Dayna R. Touron

Jordano, M. L. & Touron, D. R. (2018). How often are thoughts metacognitive? Findings from research on self-regulated learning, think-aloud protocols, and mind-wandering. *Psychonomic Bulletin & Review, 25(4),* 1269-1286.

This is a post-peer-review, pre-copyedit version of an article published in *Psychonomic Bulletin & Review*. The final authenticated version is available online at: https://doi.org/10.3758/s13423-018-1490-1

Abstract:

Metacognitive monitoring refers to how people evaluate their cognitive performance. An extensive literature examines *how accurately* individuals engage in monitoring. The question of *how often* individuals engage in metacognitive monitoring has been largely neglected, although one might expect situational, group, and individual variability in monitoring frequency. We argue that this is a critical omission, given that the frequency of metacognitive monitoring might have important implications for monitoring accuracy and task performance. Within this review, we highlight findings from three literatures, that each provide insight into how often individuals engage in monitoring. To clarify the important links and potential overlaps between these separate bodies of research, we begin by summarizing the metacognitive monitoring literature, including age-related patterns in monitoring accuracy. We then connect these questions regarding spontaneous monitoring, including age-related patterns in spontaneous monitoring, to targeted reviews of the self-regulated learning, think-aloud protocol, and mind-wandering literatures. We discuss situational and dispositional factors believed to influence monitoring accuracy, and propose that the same factors could potentially influence the frequency of spontaneous monitoring. Additionally, we propose that age-related *increases* in spontaneous monitoring (as suggested by age-related increases in TRI) may contribute to older adults' intact monitoring abilities. It is our hope that this review will encourage increased attention and research on the topic of spontaneous monitoring.

Keywords: metacognition | mind wandering | cognitive aging | monitoring

Article:

Introduction

Imagine you are sitting down to read a lengthy and complicated scientific article. It is important that you understand and can later remember the information it contains. As you read through the dense article, you find yourself engaging in extraneous thoughts. You might realize you are getting distracted by noises in your environment and stop reading until you relocate to a quieter place. Or perhaps you periodically quiz yourself on what you had just read. If you find your comprehension lacking, you may re-read a critical section. These are just some examples of *metacognition*, or thoughts one has regarding their own cognitive processes (Flavell, 1979).

Three distinct yet inter-related aspects of metacognition include knowledge, control, and monitoring. *Metacognitive knowledge* refers to a person's understanding and beliefs about cognition. *Metacognitive control* refers to regulating one's ongoing cognitive activities. *Metacognitive monitoring* refers to various thoughts people have regarding their ongoing task performance or progress. Metacognitive thoughts about one's task knowledge, performance, and approach are expected to occur with varying frequency in everyday life, but the question of *how frequently* individuals engage in metacognitive thoughts has been largely neglected within the study of metacognition.

In this review, we discuss the importance of studying the frequency of individuals' metacognitive thoughts, with the goal of increasing attention and future research on the issue. The specific focus of this paper considers the frequency of *metacognitive monitoring*, or thoughts regarding task performance. An extensive literature has examined metacognitive monitoring, generally focusing on whether individuals' judgments about their task performance accurately reflect how well they are actually performing, and which factors predict or influence the accuracy of one's performance-related thoughts. The question of *how frequently* individuals spontaneously monitor their task performance is an also critical but comparatively neglected research question. It is possible that the degree to which one spontaneously reflects on their task performance influences the accuracy of those thoughts and, ultimately, performance on the associated cognitive task.

The question of *age-related patterns* in spontaneous thoughts about task performance is also of particular interest. Older adults' metacognitive monitoring abilities remain relatively spared despite declines in other types of cognitive functioning (Dunlosky & Connor, 1997; Hasher, Zacks, & May, 1999; Salthouse, 1991; Shaw & Craik, 1989) but no work within the metacognition literature has examined the specific *frequency* and *content* of monitoring experiences in older adults. We argue that age-related patterns in the frequency or content of spontaneous metacognitive thoughts may influence age-related patterns in monitoring accuracy and task performance.

In discussing the importance of studying the frequency of individuals' thoughts regarding task performance, we begin with a brief overview of metacognitive monitoring, including discussion of age-related patterns in monitoring and other influences on metacognitive monitoring accuracy. Following this review of the monitoring literature, we consider how findings from the self-regulated learning, think-aloud protocol, and mind-wandering studies provide an indication of how frequently individuals think about their task performance. Finally, we present avenues for future research that focus on the frequency of spontaneous regarding task performance in both younger and older adults.

The relationships between cognition and different types of metacognitive thoughts can be understood by examining popular models of metacognition, such as the Nelson and Naren model (1990; see Fig. 1) and the COPES model of self-regulated learning (Winne & Hadwin, 1998). These models include an interplay between one's ability to use metacognitive monitoring to evaluate task performance and one's ability to use the results of that monitoring to alter task approach and thus improve task performance. Although these models assume that people reflect on their task performance spontaneously, few studies address the question of how often individuals *spontaneously* engage in metacognitive thoughts, or how frequently people engage in metacognitive thoughts that are *unprompted* by task instructions that ask them to reflect on their task performance.

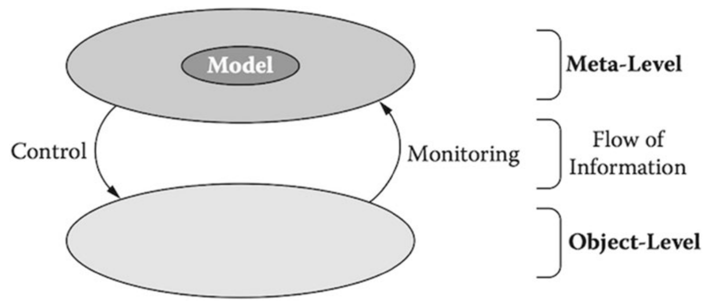

Fig. 1. The Nelson and Narens (1990) framework demonstrating the relationship between metacognition ("meta-level") and cognition ("object-level"), showing how metacognitive monitoring is used to check and update one's model of an ongoing task

Determining the accuracy of peoples' insights regarding their memory processes involves comparing ones' performance judgments to their actual task memory performance. Monitoring accuracy is a topic of great interest because accurate monitoring should lead to more updating of one's task representations, more effective regulation of task activities and, ultimately, improved task performance. Monitoring and control processes are believed to operate in a feedback loop ("monitoring affects control hypothesis"; Nelson & Leonesio, 1988). If peoples' thoughts regarding their task performance are accurate then effective strategies can be chosen to complete the task and improve performance. Individuals with better memory performance (Maki & Berry, 1984) and study strategies (Connor, Dunlosky, & Hertzog, 1997) have been shown to have better monitoring accuracy.

Examples of metacognitive judgments include *judgments of learning* (JOLs; Arbuckle & Cuddy, 1969), *feeling of knowing* judgments (FOKs; Hart, 1965), and *confidence judgments* (CJs). JOLs are made during encoding and tap into individuals' thoughts about how well they have learned an item (Arbuckle & Cuddy, 1969). FOKs are made during retention or retrieval of information and are judgments regarding the likelihood of recognizing currently unrecallable answers on a later test (Hart, 1965; Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982). CJs are made after a participant has retrieved an answer and reflect participants' thoughts about how well they were able to recall information.

Factors that influence monitoring accuracy

In the following sections, we briefly discuss *task-related* predictors and *individual-level* characteristics that may influence how well people evaluate their task performance. While these factors have been shown to influence monitoring accuracy, they may also influence the degree to which individuals spontaneously think about their task performance.

Situational predictors of monitoring accuracy

Individuals consider different and multiple cues when they think about their task performance and make metacognitive judgments (Hertzog, Hines, & Touron, 2013; Kelemen, Frost, & Weaver, 2000). *Cues* refer to information that individuals consider when thinking about task performance (Brunswick, 1943; Koriat, 1997; Nelson, 1996), such as processing fluency or item difficulty (Dunlosky & Metcalfe, 2009). *Cue utilization* refers to whether individuals use cues when making their metacognitive judgments. Cues are considered diagnostic if they predict criterion performance and non-diagnostic of they do not predict criterion performance. Participants often do not properly use diagnostic cues to make accurate metacognitive judgments, and instead use non-diagnostic cues (Hertzog et al., 2002).

Older adults and younger adults seem to be influenced by similar types of cues and there is little change in JOL resolution with increased age (Dunlosky & Hertzog, 2000; Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Hertzog, Sinclair, & Dunlosky, 2010; Robinson et al., 2006). However, the number of cues used at a given time depends in part on one's available cognitive resources (Griffin, Wiley, & Thiede, 2008). Given that executive control abilities decrease with age, older adults are less likely to use multiple cues when making metacognitive judgments (Ferguson, Hashtroudi, & Johnson, 1992). Cues used to make metacognitive judgments may also influence one's overall *propensity* to engage in monitoring; this possibility is discussed in more detail later in the paper.

Individual-level influences of monitoring accuracy

Characteristics of the individual making the judgments also influence monitoring accuracy. Metacognitive processes (including monitoring) can increase cognitive load (Coutinho, Redford, Church, Zakrzewski, Couchman, & Smith, 2015; Stine-Morrow, Shake, Miles, & Noh, 2006), so it is unsurprising that executive control ability has been found to influence one's ability to think about task performance. At least in younger adults, studies have found a positive association between working memory capacity and monitoring accuracy (Kelemen, Frost, & Weaver, 2000; Komori; 2016).

Normal cognitive aging is characterized by declines in a wide variety of cognitive tasks. Cognitive aging is associated with poorer performance on memory tasks (Lovelace & Marsh, 1985; Rabinowitz, Ackerman, Craik, & Hinchley, 1982; Shaw & Craik, 1989), declines in processing speed (Salthouse, 1991), executive control (Hasher, Zacks, & May, 1999; McDowd & Craik, 1988), and rate of learning (Dunlosky & Connor, 1997). Given these age-related cognitive declines, one might expect that older adults would also suffer relative to younger adults in monitoring ability. However, most evidence suggests age invariance in monitoring accuracy.

Younger adults have above-chance relative JOL accuracy (Hertzog, Kidder, Powell, Moman, & Dunlosky, 2002; Leonesio & Nelson, 1990; Robinson, Hertzog, & Dunlosky, 2006), and older adults have similar or higher relative JOL accuracy, at least when older adults are asked to provide item-by-item (Baker, Dunlosky, & Hertzog, 2010; Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Hertzog & Dunlosky, 2011; Hertzog & Hultsch, 2000; Hertzog, Kiddler, Powell-Momon, & Dunlosky, 2002; Hertzog, Sinclair, & Dunlosky, 2010) rather than global predictions of performance (Brigham & Pressley, 1988; Hertzog, Price, & Dunlosky, 2008; Hertzog, Touron, & Hines, 2007). Absolute JOL accuracy may also be equivalent in younger and older adults (McDonald-Miszczak, Hunter, & Hultsch, 1994).

As with JOLs, younger adults' FOKs are moderately correlated with task performance (Hart, 1967; Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982) and absolute accuracy for FOK judgments about general knowledge also shows age invariance (Butterflied, Nelson, & Peck, 1988; Eakin, Hertzog, & Harris, 2014; Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007). Findings on age-related differences in the accuracy of CJs have been mixed; some work suggests that younger and older adults are equally good at making post-dictions regarding memory performance (Hertzog, Saylor, Fleece, & Dixon, 1994), whereas other studies suggest that older adults are more likely than younger adults to produce high-confidence false alarms (e.g., Dodson, Bawa, & Krueger, 2007) and show poorer resolution in CJs when incorrect lures are present (e.g., Kelley & Sahakyan, 2003), indicating that older adults may poorly encode information (Hertzog & Dunlosky, 2011).

Although there is evidence that monitoring accuracy is largely spared with age, participants' expectations and beliefs regarding their memory ability (Crumley, Stetler, & Horhota, 2014) can influence monitoring accuracy. For example, as people age the association between their subjective and objective memory performance increases (Crumley, Stetler, & Horhota, 2014). Subjective memory is typically measured by having participants self-rate their general memory ability, which is then compared to objective performance on a memory test. As participants age they tend to do worse on objective memory tasks but they also report worse subjective memory performance (Cook & Marsiske, 2006; Crumley, Stetler, & Horhota, 2014; Hülür, Hertzog, Pearman, & Gerstorf, 2015; Wahlin, Maitland, Backman, & Dixon, 2003; Zelinski, Gilewski, & Thompson, 1980). Older adults may become increasingly aware of their memory deficits as they age.

Memory self-efficacy also influences the thoughts individuals have regarding their task performance. Older adults tend to report lower memory self-efficacy than younger adults (Cavanaugh & Poon, 1989; Hertzog et al., 1990; Hultsch, Hertzog, & Dixon, 1987). Older adults with higher memory self-efficacy not only have better memory performance, but they also report higher predictions of their memory ability (Hertzog et al., 1990). Therefore, one's global memory self-efficacy appears to influence one's subjective memory and monitoring ability.

Differences in personality are also associated with monitoring accuracy. Participants who score highly on the characteristics of *openness to new experiences* and *extraversion* tend to have higher confidence in their JOLs than those who score lower on these characteristics (Buratti, Allwood, & Kleitman, 2013). Relationships between extraversion and overconfidence (Dahl, Allwood, Rennemark, & Hagberg, 2010; Pallier, Wilkinson, Danthir, Kleitman, Knezevic, Stankov, &

Roberts, 2002; Schaefer, Williams, Goodie, & Campbell, 2004) and openness and overconfidence (Dahl et al., 2010; Schaefer, Williams, Goodie, & Campbell, 2004) have also been observed. Older adults high in agreeableness and consciousness have stronger associations between their subjective memory ratings and objective memory performance, suggesting that older adults high in these personality factors monitor more accurately than those who are low in these factors (Hülür, Hertzog, Pearman, & Gerstorf, 2015).

Factors that influence how accurately people monitor, such as cognitive resources and personality characteristics, may also contribute to individual differences in the frequency of metacognitive thoughts one experiences while completing a cognitive task. This possibility is discussed in more detail later in the paper.

Assessing monitoring frequency

As mentioned, studying the accuracy of metacognitive thoughts requires participants to make explicit judgments (i.e., JOLs, FOKs, and CJs) about their cognitive abilities. Research examining the accuracy of these metacognitive judgements provide us with information regarding how accurately people monitor, but they tell us less about how often they monitor. Although studies focusing on the accuracy of metacognitive judgments may not be useful in answering questions regarding monitoring frequency, other methodologies can be used to determine often people monitor, age-related changes in monitoring frequency, and situations and characteristics that alter propensity to engage in monitoring. First, we will discuss studies from the metacognition literature examining self-regulated learning and spontaneous strategy use, given that individual-differences in spontaneous use of certain learning strategies may reflect differences in spontaneous metacognitive monitoring. Secondly, we will discuss metacognition studies conducted using think-aloud protocols. Finally, we will discuss studies from the mind-wandering literature that focus on frequency of off-task thoughts about metacognition.

Findings from self-regulated learning

As described above, a substantial literature suggests that monitoring ability is robust and remains fairly intact with age. However, there is little work addressing how *frequently* individuals spontaneously think about their task performance. One area of research that can address the question of monitoring frequency focuses on self-regulated learning. In these studies, participants are asked to study information for a later test, and participants' allocation of study time across items can be examined for the extent to which learning and later test performance are optimized. As described in the Nelson and Narens (1990) model presented in Fig. 1, metacognition involves both monitoring and control processes, where cognitive activity (such as learning) is represented at the meta-level via monitoring processes, and this representation may inform or produce actions (such as the regulation of study time) via control processes. Spontaneous self-regulation of study is a classic case of the interplay between these two components of metacognition.

Research in this area often focuses on individual differences in self-regulated learning and strategy use. For example, some individuals more efficiently allocate their study time. Academically successful fifth-grade students show sensitivity to the difficulty of to-be-studied texts (Owings, Petersen, Bransford, Morris, & Stein, 1980), spontaneously allocating more study time towards texts rated as difficult compared to texts rated as easy. In contrast, less successful students studied equally for both difficulty and easy texts, whereas these students could distinguish text difficulty and adjust their allocation of study time when prompted, they did not *spontaneously* do so. The ability to spontaneously allocate one's study time improves across childhood, with younger children (grades 1 and 3) spontaneously engaging in less efficient use of study time than older children (grades 5 and 7) (Dufresne & Kobasigawa, 1989). Self-paced study in college students shows similar findings. More successful college students study harder items longer than easy items, whereas less successful college students do not spontaneously allocate more study time towards studying difficult items (Cull & Zechmeister, 1994).

Differences between younger and older adults in self-regulated learning have also been examined, with mixed findings (see Bottiroli, Dunlosky, Guerini, Cavallini, & Hertzog, 2010). In one study, older and younger adults were asked to think aloud when studying sets of pictures matched in difficulty for a later serial recall test. Younger adults spontaneously studied longer, rehearsed the information more, and had better eventual recall than older adults for difficult items (Murphy, Sanders, Gabriesheski, & Schmitt, 1981). Older adults did not study the difficult items or rehearse the items unless specifically instructed to do so. In this aging study and in the previously mentioned studies of children and college students, spontaneous allocation of study time towards difficult items is believed to reflect more spontaneous monitoring, given that both children and younger adults can allocate study time to more difficult items but do not readily do so unless specifically prompted to. Older adults have also been found not to study difficult items for as long as younger adults do, unless they are specifically given instructions that encourage increased monitoring (Murphy, Schmitt, Caruso, & Sanders, 1987).

These findings suggest that some people spontaneously engage in monitoring of task features such as stimuli difficulty. Individuals who more effectively monitor and self-regulate their learning strategies achieve better task performance than those who presumably do not as readily engage in monitoring. While the self-regulated learning work suggests variability within and between age groups in spontaneous monitoring and control as well as age-related declines in these processes, it is important to note that study time allocation most directly reflects metacognitive control rather than monitoring. The specific frequency of spontaneous monitoring and the content of monitoring processes can be inferred given the relationships between monitoring and control, but are not measured in these studies. More detailed information regarding individual differences in monitoring frequency and the frequency of different types of monitoring can be found by examining the think-aloud literature.

Findings from think-aloud protocols

Studies using *think-aloud protocols* (Ericsson & Simon, 1993; Veenman, Elshout, & Groen, 1993) can provide insight into how often individuals engage in spontaneous monitoring. Using think-aloud protocols, the specific content and frequency of individuals' metacognitive monitoring can be reported and measured. Within this section, we provide a brief overview of the think-aloud protocol approach along with a discussion of concerns regarding their veridicality and reactivity. Next, we briefly discuss the strengths and limitations associated with using think-aloud protocols in examining metacognition. Finally, we discuss findings from the think-aloud literature that provide insight into how often children, younger adults, and older adults engage in spontaneous monitoring.

Think-aloud methodology

Think-aloud protocols (Ericsson & Simon, 1984) have been widely used to investigate and build models of cognitive and metacognitive processes. In this research, participants (also known as "raters") are instructed to verbalize their thoughts while completing a cognitive task. Think-aloud protocols have been used to study reading comprehension, writing, mathematical problem-solving, and other types of problem-solving. Rater's verbalizations of their thought processes are recorded, transcribed, and then analyzed to identify the metacognitive and decision-making processes that are used when completing these tasks. Due to the rich nature of these data, metacognitive thoughts can be distinguished in fine detail, including whether they involve monitoring or control processes.

Think-aloud protocols can garner large amounts of information regarding specific behaviors and thought processes, but there are criticisms and limitations to this approach. Because raters are typically not accustomed to verbalizing their internal thought processes while simultaneously focusing on completing a task (Smagorinsky, 1994), think-aloud protocols can be difficult to administer. Participants must receive instructions regarding how to verbalize their thoughts and often need to be prompted to continue verbalizing during the task. Additionally, gathering and scoring data from individual raters' think-aloud protocols is time-consuming and experimenters must be trained on how to properly score think-aloud protocols. Therefore, using think-aloud protocols can be challenging with large participant samples (Azevedo et al. 2010; Schellings & Broekkamp, 2011).

Other perhaps more fundamental criticisms of think-aloud protocols involve their veridicality and reactivity. Veridicality concerns whether think-aloud protocols accurately represent raters' actual and complete thinking. Reactivity concerns whether think-aloud protocols alter either the processes being observed or the outcomes of those processes (i.e., task performance). Ericsson and Simon (1984) note that veridicality depends on the nature of the task being completed. Think-aloud protocols are believed to be most accurate for the contents of raters' short-term memory, which should be accessible for conscious introspection as long as rater verbalizations are concurrent to the task of interest (Ericsson & Simon, 1987; Lumley, 2005).

Regarding the reactivity of think-aloud protocols, Ericsson and Simon (1984, 1993) argue that this depends on the type of verbalization requested. It is advised that raters merely verbalize their thought processes and behaviors without interpreting them (Ericsson & Simon, 1984, 1993). A meta-analysis (Russo, Johnson, & Stephens, 1989) indicated that think-aloud protocols result in significantly longer response times (Fidler, 1983; Karpf, 1973) and fewer problems being solved within a fixed time period (Fryer, 1941). However, think-aloud protocols are not associated with changes in task accuracy (Leow & Morgan-Short, 2004; Barkaoui, 2011) and are thus not believed to alter the underlying thought processes and behaviors associated with the task.

Think-aloud protocols and monitoring frequency

Think-aloud studies sometimes include the exact frequencies of verbalized metacognitive thoughts, which provides an indication of how frequently people spontaneously think about task performance. As with the self-regulated learning literature, findings from the think-aloud protocol literature suggest that some individuals more spontaneously engage in metacognitive monitoring than others and that those who do chose more effective strategies and end up obtaining better task performance. Think-aloud protocols have been used to examine differences in metacognition between high-achieving and low-achieving students and between expert and non-expert problem solvers.

For example, think-aloud protocols collected during a reading comprehension task reveal that fourth and fifth grade students vary considerably in how frequently they reported monitoring (Meyers, Lytle, Palladino, & Green, 1990). Across all students, verbalizations that reflected metacognitive monitoring made up 8.5% of the total number of verbalizations, with a range of between 0% and 27% of students' verbalizations during the task. While the number of monitoring verbalizations as a percentage of total verbalizations was not broken down by reading comprehension scores in this study, Meyers et al. (1990) did note that the more successful readers verbalized monitoring more frequently than unsuccessful readers. The finding that children who are successful readers verbalize more monitoring has been corroborated by other studies (Garner & Taylor, 1982; Janssen et al., 2006; Yuill & Oakhill, 1991).

Studies of monitoring frequency using think-aloud protocols for reading tasks with older children and young adults have found comparable results. "Adequate readers" aged 12–18 years report more monitoring for both informational and narrative texts compared to "poor readers" of the same age (Denton, Enos, York, Francis, Barnes, Kulesz, Fletcher, & Carter, 2015). Among high-school students, more academically successful students report more monitoring during reading tasks than less academically successful students (Janssen et al., 2006). These findings all suggest that some children and young adults engage in more spontaneous monitoring than others, and these increases in spontaneous monitoring may influence task approach and task performance.

The type of reading material being studied has also been found to alter monitoring frequency. In studies using think-aloud protocols during reading tasks, college students who read *narrative texts* were found to make fewer metacognitive comments than students who read informative texts (Trabasso & Magliano, 1996) and college students who read narrative texts recreationally reported less monitoring than students who read informative texts (Narvaez et al., 1999).

Think-aloud protocols have also been used to examine differences in monitoring during mathematical problem-solving. Gifted students verbalize monitoring more often than both average-achieving and learning-disabled children (Montague & Applegate, 1993). Additionally, students verbalize monitoring more when they are completing open-ended math problems and complex math problems compared to easier problems (Mokos & Kafoussi, 2013). Various think-aloud protocol studies of other types of problem-solving have found that expert problem-solvers more spontaneously use self-monitoring to improve task performance (for a review, see Glaser, 1991). As with reading comprehension tasks, there is individual variability in spontaneous monitoring, with those who report more spontaneous monitoring also obtaining better task performance.

Schellings et al. (2013) used think-aloud protocols to assess the frequency of different types of metacognitive activities. In this study, protocols for 20 high school students were recorded as they read a history text for an upcoming exam. On average, 30.6% of the reported activities were examples of metacognitive monitoring (Schellings, Hout-Wolters, Veenman, & Meijer, 2013), with participants varying considerably in how much monitoring they reported. Other work that has used think-aloud protocols to examine the frequency of various metacognitive thoughts during learning of a foreign language (Leow & Morgan-Short) and reading comprehension tasks (Martin, 1987) has found similar frequencies of metacognitive monitoring.

Instructions to engage in self-testing have been found to increase monitoring. Fernandez and Jamet (2017) found that college students spontaneously report monitoring task performance on roughly 28% of their verbalizations and spontaneously report monitoring the effectiveness of their task strategies on roughly 5% of their verbalizations. When given instructed to engage in self-testing, verbalizations about performance monitoring increased to roughly 49% and verbalizations about strategy monitoring increased to roughly 14%. Additionally, performance was improved when raters were asked to engage in self-testing. Spontaneous monitoring frequency is also associated with strategy use in regular and in computer-learning environments (Deekens, Greene, & Lobczowski, 2018).

Think-aloud protocols, monitoring frequency, and age

In addition to being used in studied with children and younger adults, think-aloud protocols have been used to study metacognition in older adults and to examine age-related changes in metacognition. One study comparing think-aloud protocols obtained during the Tower of London Task found that the number of verbalizations about monitoring did not differ between younger and older adults (Gilhooly, Phillips, Wynn, Logie, & Della Sala, 1999). This finding suggests age-invariance in monitoring frequency. While few think-aloud protocol and aging studies have directly examined frequency of metacognitive verbalizations, some additional think-aloud protocol research has found that younger and older adults report equivalent generation of mediators when studying lists of words, which could indicate that younger and older adults both engaged in spontaneous monitoring during the task (Fox, Baldock, Freeman, & Berry, 2016).

While studies such as these provide an idea of how often spontaneous monitoring occurs, and suggest that there are individual differences in propensity to spontaneously monitor, many think-aloud studies do not report the specific frequencies of verbalizations corresponding to monitoring. As such, this literature does not currently provide a clear indication of exactly how often frequently monitoring occurs, how much individuals vary in their frequency of monitoring, or how monitoring frequency changes across the lifespan.

Findings from mind-wandering research

The mind-wandering literature may provide additional insight into how often individuals spontaneous engage in metacognitive thoughts. It is common for one's attention to become decoupled from the task they are completing and instead become directed towards internally generated thoughts, an experience commonly referred to as *mind-wandering* (Giambra, 1995, Smallwood & Schooler, 2006). As with concurrent think-aloud protocols, studies of mind-wandering often require participants to report the types of thoughts they experience as they perform a cognitive task. While the think-aloud protocol literature focuses on the types of on-task and metacognitive thoughts individuals experience during completion of cognitive tasks, the mind-wandering literature focuses on the frequency of both task-relevant and task-irrelevant thinking that occurs during task completion.

An early theory of mind-wandering, *current-concerns theory* (Klinger, 1971), originated to explain the flow of conscious, internally-generated thoughts. This theory postulates that, once a person establishes a particular goal, this goal becomes a *current concern.* Current concerns are kept accessible until they are either abandoned or resolved, and can therefore be cued by stimuli in one's environment as well as by their other thoughts (Klinger, 1971). This can result in individuals experiencing different types of off-task thoughts, depending on the current concerns experienced within a given environment.

Individuals have a wide variety of thoughts throughout the course of a day, which can be conceptualized along two dimensions: "task-relatedness" and "stimulus-dependency" (Stawarczyk, Majerus, Maj, van der Linden, & D'Argembeau, 2011a). Experiencing thoughts that are task-related and stimulus-dependent indicate that one is completely focused on performing the task. Intrusive thoughts that are stimulus-dependent but task-unrelated include external sources of distraction (Forster & Lavie, 2008; Unsworth, Redick, Lakey, & Young, 2010). Intrusive thoughts that are stimulus-independent and task-related include interfering thoughts that are related to task appraisal and include metacognitive thoughts (Matthews, Joyner, Gililand, Campbell, & Falconner, 1999; Smallwood, Davies, Heim, Finnigan, Sudberry, O'Connor, & Obonsawin, 2004). Finally, thoughts that are both stimulus-independent and task-unrelated include mind-wandering about things that are *completely unrelated* to the task and environment (Antrobus, Singer, Goldstein, & Fortgang, 1970; Giambra, 1989; Smallwood & Schooler, 2006; Teasdale, Dritschel, Taylor, Prcotor, Llyod, Nimmo-Smith, & Baddeley, 1995).

The frequency of these categories of thoughts can be assessed in several ways. Mind-wandering can be studied using retrospective questionnaires (Lachman & Agrigoroaei, 2012; Maillet & Rajah, 2013). However, as with retrospective think-aloud assessments, using retrospective thought content questionnaires carries the risk of participants either forgetting or incorrectly reconstructing the thoughts experienced during the task of interest. This is particularly problematic for older adults who experience both memory declines and difficulty in aggregating information regarding their cognitive performance (Hertzog et al., 2008; Hertzog, Touron, & Hines, 2007).

Mind-wandering can instead be studied by having participants engage in a cognitive task during which their intrusive thoughts are either "self-caught" or "probe-caught." With the "self-caught" method, participants must monitor their thoughts during the task and indicate when they notice their mind is wandering from the task. In studies using the "probe-caught" method, participants instead respond to thought probes that periodically interrupt the task. These probes ask participants to indicate the type of thought that they were experiencing right before the probe appeared. Because the self-caught method requires participants to monitor their internally-generated thought processes and is thus more cognitively demanding, the probe-caught method is more commonly used, especially in studies of aging and mind-wandering.

Probes vary in the questions they include, but often ask participants to classify the type of thought they experienced *immediately prior* to the appearance of the probe. Thought probes are typically positioned approximately 2 to 3 min apart to allow the mind to begin wandering across task trials (see Seli, Carriere, Levene, & Smilek, 2013). An example of a commonly-used thought probe appears below. As participants complete computerized cognitive tasks, they are occasionally interrupted and asked "*What were you just thinking about*?" Participants select one of the following options using their computer keyboard:

(1) The task: Focused on completing the task, verifying equations and remembering letters
(2) Task experience/performance: Evaluating one's performance
(3) Everyday things: Thinking about recent or impending life events
(4) Current state of being: Thinking about conditions such as hunger or sleepiness
(5) Personal worries: Thinking about concerns, troubles or fear not relating to the experimental task
(6) Daydreams: Fantasies disconnected from reality
(7) Other

The response options correspond to different types of thoughts the participants may have been experiencing. Here, option (1) corresponds to being completely on-task and thinking only about responding to the task stimuli, option (2) corresponds to thoughts regarding task performance, and options (3) through (7) correspond to thoughts that are completely unrelated to the cognitive task at hand. Several online thought probes are embedded throughout an experimental task, and participants' responses are used to obtain mean proportions of on-task thoughts, thoughts about things unrelated to the task, and thoughts about task performance or strategy.

Mind-wandering reports obtained using probes such as the one above seem to provide an accurate reflection of participants' underlying thought processes. Participants' probe-caught mind-wandering is associated with higher in-the-moment task errors relative to on-task reports (McVay & Kane, 2012; McVay, Meier, Touron, & Kane, 2013) and is also correlated with measures of attentional lapses such as intraindividual response time variability (McVay & Kane, 2009; McVay & Kane, 2012; McVay, Meier, Touron, & Kane, 2013). Importantly, probe-caught thought reports do not simply reflect participants' reactive explanations for their own performance. In other words, participants do not seem to assume that mind-wandering has occurred whenever they make performance errors. Research comparing different types of errors find that mind-wandering is associated only with target misses but not lure false alarms, supporting the notion that participants' probe-caught thought reports are a valid measure of their underlying thought processes (McVay, Meier, Touron, & Kane, 2013). Additionally, in a study examining the effects of mind-wandering on younger and older adults' eye-movements and reading comprehension (Frank, Nara, Zavagnin, Touron, & Kane, 2015) individuals in both age groups made more regressions and engaged in more blinking when reporting mind-wandering than when they reported being on-task (Frank et al., 2015), further validating that mind-wandering reports accurately reflect attentional lapses.

Task-related interference

Particularly relevant to the question of how often individuals spontaneously monitor their task performance is the smaller body of literature examining intrusive thoughts that are both stimulus-independent and task-related. Such thoughts are referred to as "task-related interference" (TRI). Because TRI encompasses thoughts regarding task performance and task appraisal, TRI appears to be analogous to metacognitive monitoring. As such, and critical to the current review, studies examining the frequency and content of TRI can provide insight into how and how frequently individuals monitor their task performance.

Even though TRI encompasses metacognitive thoughts, within the mind-wandering literature TRI is often characterized as a type of mind-wandering. Both TUTs and TRI have been linked to in-the-moment performance deficits (McVay & Kane, 2009; McVay et al., 2013; Mrazek, Chin, Schamder, Hartson, Smallwood, & Schooler, 2011; Smallwood, O'Connor, & Heim, 2005), indicating that, while TRI experiences are stimulus-dependent, they are distinguishable from on-task thoughts where participants are entirely focused on responding appropriately to task demands.

Factors that influence TRI

Early research focusing on thoughts about task appraisal was conducted in the testing anxiety and cognitive interference body of literature (Peterson, Swing, Braverman, & Buss, 1982; Sarason, Sarason, Keefe, Hayes, & Shearin, 1986). This early research examining off-task thoughts about task appraisal suggests that some individuals are more likely than others to experience task-related interference (Sarason, Sarason, Keefe, Hayes, & Shearin, 1986). Various factors influence the amount of TRI (metacognitive monitoring) experienced while completing a cognitive task (Frank et al., 2015; Jordano & Touron, 2017a; Jordano & Touron, 2017b; McVay et al., 2013; Smallwood et al., 2009).

For example, increased negative affect is associated with increased reporting of both TUTs (Killingsworth & Gilbert, 2010; Smallwood et al., 2009) and TRI (Smallwood et al., 2009). Task difficulty is also associated with propensity to engage in mind-wandering. McVay et al. (2013) found that frequency or TUTs and TRI were differentially affected by task difficulty. Younger adults reported fewer TUTs during a more challenging task than during an easier task, perhaps because they had fewer spare cognitive resources available to engage in TUTs. In contrast, both younger and older adults report more TRI during challenging tasks compared easier tasks, indicating that participants have more evaluative thoughts about their performance and engage in more monitoring when the task is difficult. Similarly, Zavagnin et al. (2014) found that that older adults report more TRI during difficult tasks compared to easier tasks. The results of these studies suggest that manipulating task difficulty can increase one's propensity to engage in metacognitive monitoring.

Age influences both the amount of probe-caught mind-wandering and the content of those mind-wandering experiences. While they report fewer TUTs and less overall mind-wandering compared to younger adults, older adults do report more TRI than younger adults (Frank et al., 2015; Jordano & Touron, 2017a; McVay et al., 2013; Zavagnin, Borella, & De Beni, 2014).

This may be surprising given age-related declines in executive control (Hasher, Zacks, & May, 1999). A framework of mind-wandering by McVay and Kane (2010) can accommodate these age-related patterns in mind-wandering. The "Control failures × Current concerns" account posits that intrusive thoughts are automatically and continuously generated in response to various environmental cues that prime an individual's current goals and concerns (see Klinger 1978; Klinger & Cox, 1987-1988) and that off-task thoughts *only* enter conscious awareness through a failure of executive control resources to suppress those thoughts. Therefore, individual differences in mind-wandering are believed to involve an interaction between one's executive control capabilities, motivation to prevent off-task thoughts from entering consciousness, *and the extent to which environmental context primes current concerns* (McVay & Kane, 2010). From this perspective, older adults have fewer TUTs and more TRI compared to younger adults in part because their current concerns differ from those of young adults. A typical laboratory testing environment on campus may trigger young adults' current concerns regarding their everyday school-related concerns, leading to increased TUTs. This same testing environment may not cue older adults' typical, everyday current concerns (Carstensen, 1995), instead triggering concerns about age-related cognitive decline (Hertzog & Hultsch, 2000), leading to increased TRI.

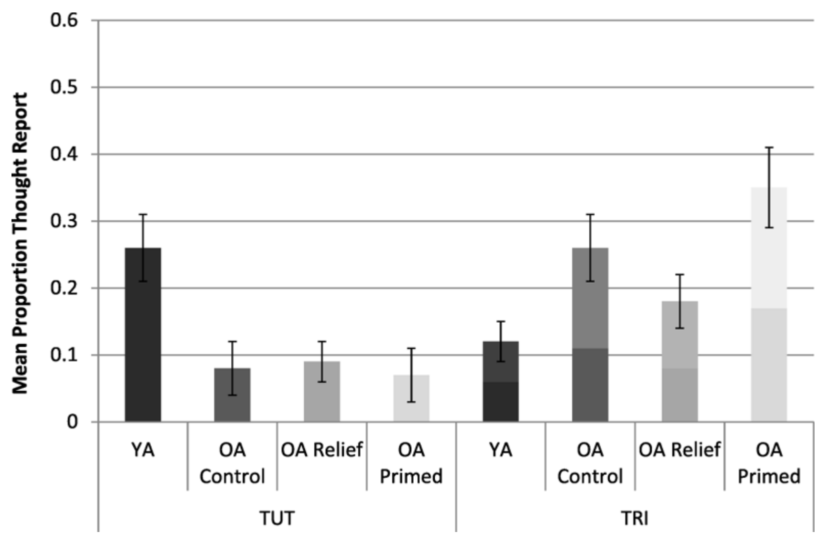

Fig. 2. Mean proportions of probe-caught thought types during the OSPAN (Jordano & Touron, 2017). Bars indicate *SE*s. TRI is broken down into proactive TRI (thinking about task strategy or approach) represented by the lighter portions of the TRI columns and reactive TRI (thinking about task performance) represented by the darker portion of the TRI columns. *SE* bars on the TRI columns represent the *SE* for overall mean proportion of TR. *TRI* task-related interference, *TUT* , *YA* younger adults, *OA* older adults, *TR*

Some studies have manipulated the salience of current concerns and shown an impact on the type and amount of mind-wandering that people experience. Priming people to think about their unfulfilled, everyday goals leads to increased reporting of TUT experiences (Kopp, D'Mello, & Mills, 2015; Masicampo & Baumeister, 2011; McVay & Kane, 2013). Likewise, the priming of personal, *performance-related* concerns increases the amount of TRI younger (Jordano & Touron, 2017b) and older adults (Jordano & Touron, 2017a; see Fig. 2) report. The salience of performance-related concerns can be manipulated by inducing stereotype threat (Steele & Aronson, 1995). Models of stereotype threat assume that reductions in task performance occur because secondary processes such as emotion regulation and metacognitive monitoring deplete individuals' executive control resources (Schmader, Johns, & Forbes, 2008). Given that stereotype threat manipulations increase the amount of TRI, factors believed to moderate stereotype threat may also affect the frequency of spontaneous monitoring.

Concerns about one's cognitive abilities can influence beliefs regarding the control one has over cognitive performance and cognitive change. Those with higher *control beliefs* believe that they can engage in behaviors that will allow them to maintain or even improve their cognitive functioning while getting older. Older adults with higher control beliefs have better memory performance (Lachman, Neupert, & Agrigoroaei, 2011) than those with lower control beliefs. Relationships also exist between control beliefs and mind-wandering; lower levels of perceived control negatively influence episodic memory performance by increasing retrospectively-caught TRI in both younger and older adults (Lachman & Agrigoroaei, 2012). Lower levels of perceived control may particularly increase negatively-valenced TRI.

Mindfulness, or the ability to attend to the present moment in a non-judgmental way, is believed to affect attention. Although mind-wandering and mindfulness appear to be opposing constructs, the relationship between mindfulness and mind-wandering remains tentative. A negative association between trait mindfulness and probe-caught TUTs during sustained attention tasks (Fountain-Zaragoza, Londerée, Whitmoyer, & Prakash, 2016) has been observed. Frank et al. (2015) found that only one specific aspect of dispositional mindfulness (i.e., the tendency to observe and notice one's surroundings) mediated the relationship between TRI and age. It is possible that those that score higher on measures of dispositional or trait mindfulness not only experience fewer TUTs, but also engage in more spontaneous monitoring. More work can be done examining how state and trait mindfulness affect propensity to engage in both TUTs and TRI.

TRI and task performance

Although both TUTs and TRI are associated with in-the-moment performance errors, it seems plausible that engaging in TRI could benefit task performance. Task-related interference may be analogous to metacognitive monitoring, and successfully monitoring can lead to improved metacognitive control and task performance. While some studies have found a negative association between TRI and task performance (Coy, O'Brien, Tabaczynski, Northern, & Carels, 2011; Jordano & Touron, 2017a; Jordano & Touron, 2017b; Smallwood et al., 2004; Smallwood et al., 2009), others have found a positive association (McVay & Kane, 2012) or no association (Stawarczyk, Majerus, Maj, van der Linden, & D'Argembeau, 2011b). The relationship between TRI and performance is normally examined as the effect of TRI on in-the-moment performance or for shorter tasks. While TRI may disrupt performance in the moment, it is possible that TRI leads to downstream benefits that are only observable when longer or more complex tasks are used. Further research is needed to examine potential downstream benefits of TRI.

While the studies reviewed here include TRI as a response option, it is important to note that the exact content of individuals' TRI experiences remains largely uninvestigated. While TRI is believed to include thoughts about task performance (metacognitive monitoring), it is also possible that individuals think about implementing different task strategies to improve performance (metacognitive control). More research should be conducted examining the specific content of individuals' TRI experiences, the frequency in which individuals spontaneously engage in different examples of TRI, as well as how different types of TRI affect task performance.

Directions for future research

Frequency of spontaneous monitoring in laboratory tasks

Substantial progress in the understanding of spontaneous monitoring may be achieved with continued research using the approaches described in this review. Given that mind-wandering methods can offer detailed insight and are also relatively efficient, there may be particular motivation to pursue this approach. It is important to note that many mind-wandering studies do not include TRI as a response option and those that do typically assume that participants' TRI episodes reflect off-task thoughts regarding task appraisal. Participants are not usually asked follow-up questions regarding the specific content of their TRI experiences. Because of this, it is difficult to determine whether TRI is equivalent to metacognitive monitoring, or whether TRI also encompasses other metacognitive thoughts, such as off-task thoughts regarding metacognitive control processes. In one study that provided multiple TRI subtypes to choose from (Jordano & Touron, 2017a), younger and older adults reported thinking about both task appraisal (metacognitive monitoring) and task strategy (metacognitive control). More work is needed to examine the exact content of individuals' TRI experiences and to determine whether TRI can be thought of as equivalent to metacognitive monitoring.

Theories of age-related cognitive change offer contrasting predictions regarding whether older adults are expected to engage in more or less spontaneous thinking than younger adults. From a neuroscience perspective, spontaneous thoughts seem to involve increased activation in the default-mode network of the brain (Buckner, Andrews-Hanna, & Schacter, 2008), and research suggests less connectivity in these areas for older adults (Damoiseaux, Beckmann, Arigita, Barkhof, Scheltens, Stam, Smith, & Rombouts, 2008). Although mind-wandering studies find that older adults report more off-task thoughts regarding task appraisal (Frank et al., 2015; Jordano & Touron, 2017a; McVay & Kane, 2012; McVay et al., 2013), research on other types of spontaneous cognitive thoughts support the *reduced cognitive resources theory* of aging and reduced spontaneous cognition in older adults (Craik, 1983; Craik & Byrd, 1982). It is therefore unclear whether we should expect declines in spontaneous monitoring with age.

A next step in the study of spontaneous metacognitive monitoring should be to measure the frequency of both younger and older adults' performance-related thoughts in laboratory settings across various cognitive tasks. Online thought content probes from the mind-wandering literature can provide a methodology for allowing participants to indicate the content of their thoughts while completing cognitive tasks in the lab. Such studies can provide an indication of the degree to which individuals spontaneously think about their task performance. If findings from the TRI literature do indicate the amount of spontaneous monitoring people engage in, as is generally interpreted, the increased TRI reported by older adults may reflect increased spontaneous monitoring relative to younger adults.

Frequency of spontaneous monitoring in everyday tasks

Although mind-wandering methods can indicate how frequently people spontaneously think about their task performance, the vast majority of studies have been conducted within a laboratory setting using computerized cognitive tasks. Completing tasks in the lab can be quite different from completing tasks in everyday life. While everyday tasks are likely to be familiar, laboratory tasks typically involve novel stimuli and environments. Individuals may monitor their performance more often in laboratory settings compared to everyday tasks because of this unfamiliarity or because they are being observed and assessed. In familiar everyday tasks, participants may be less concerned with how accurately they perform because the consequences of failures may be less public and less consequential.

As noted, stereotype threat may make older adults in a laboratory setting particularly likely to worry about their task performance and think about experimenter perceptions. This could increase thoughts regarding task appraisal or task difficulty and deficits in task performance. Indeed, evidence across numerous cognitive domains demonstrates that older adults perform more comparably to younger adults in everyday cognitive tasks or with more naturalistic materials (Kempe, Kalincinski, & Memmert, 2015; Park, Hertzog, Leventhal, Morrell, Levental, Birchmore, Martin, & Bennett, 1999; Rendell & Thomson, 1999; Verhaeghen, Martin & Sedek, 2012). Given the relationship between task performance and performance monitoring, it is possible that older adults' underperformance on laboratory tasks relative to everyday tasks reflects underlying differences in the frequency of older adults' monitoring across these different environments.

Experience sampling studies of mind-wandering indicate that younger adults engage in as much off-task thinking about things unrelated to the task (TUTs) during everyday life as they do in laboratory settings (Kane et al., 2007). The results from the only published study of everyday mind-wandering in older adults (Gardner & Ascoli, 2015) indicate that older adults may engage in more TUTs during everyday life tasks than they do during laboratory tasks. However, this study did not include questions regarding TRI.

Given that individuals, particularly older adults, often perform tasks differently in everyday life compared to in the lab, an important direction for future research on the topic of spontaneous metacognition will be to examine spontaneous thoughts regarding task appraisal in everyday life, perhaps by using experience sampling methods that include thought content probes containing task-related interference as a response option.

The relationship between spontaneous monitoring and task performance

As mentioned, associations exist between one's metacognitive monitoring accuracy and one's performance on criterion tasks. Although older adults underperform relative to younger adults on many tasks despite spared monitoring (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Hertzog & Hultsch, 2000; Hertzog, Kiddler, Powell-Momon, & Dunlosky, 2002; Hertzog, Sinclair, & Dunlosky, 2010; Hertzog & Dunlosky, 2011, Butterfield, Nelson, & Peck, 1988; Rabinowitz, Ackerman, Craik, & Hinchley, 1982), individuals with more accurate thoughts regarding their task performance generally tend to have better task performance than those with less accurate thoughts regarding their task performance. Additionally, individuals who choose more efficient encoding and study strategies also tend to monitor more accurately than those who do not chose effective strategies (Hertzog et al., 2010).

Although the *quality* of individuals' thoughts about their performance seems to matter for both task performance and task strategy choice, it is also possible that the *quantity* or frequency of these thoughts could influence performance and strategy choice. Although this relationship has received little focus, there is work suggesting that monitoring frequency influences task strategy and subsequent task performance. Several notable examples of this work were described earlier, including work on children's reading behaviors and success (Owings et al., 1980), experts' processing of math problems (Schoenfeld, 1985, 1987), students' learning from history and science texts (Deekens, Greene, & Lobczowski, 2017), and the benefits of spontaneous self-testing on older adults' memory performance (Bailey, Dunlosky, & Hertzog, 2010; Bottiroli et al., 2010; Murphy et al., 1987).

Table 1. Potential task-related predictors of spontaneous monitoring

Task-related predictor	**Effect on monitoring accuracy**	**Relevant citations**	**Potential effect on monitoring frequency**
Perceptual fluency	YAs and OAs use these cues similarly. More fluent items are judged to be more likely to be remembered.	Simmons & Nelson, 2006; Tversky & Kahneman, 1974; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Castel, McCabe, & Roediger, 2007; Alter & Oppenheimer, 2006; Reder, 1987	Fluency can be manipulated within lists of items and effects on monitoring frequency examined. Increasing perceptual fluency may decrease spontaneous monitoring
Cues regarding item difficulty	YAs and OAs use these cues similarly. Difficult test items are judged as less likely to be remembered	Undorf & Erdfelder, 2013	Item difficulty can be manipulated within lists of items and effects on monitoring frequency examined. Increasing item difficulty may increase spontaneous monitoring
Item familiarity	YAs and OAs use this cue similarly. Increasing familiarity can result in participants becoming overconfident in their JOLs	Shanks & Serra, 2014	Familiarity can be manipulated within lists and effects on monitoring frequency examined. Increasing item familiarity may decrease spontaneous monitoring
Concreteness and associative relatedness	YAs and OAs use this cue similarly. Decreasing concreteness and relatedness results in items being judged as less likely to be remembered	Rabinowitz, Ackerman, Craik, & Hinchley, 1982	Concreteness and relatedness can be manipulated within lists and effects on monitoring frequency examined. Increasing concreteness and relatedness may decrease spontaneous monitoring
Memory for past performance	YAs and OAs use this cue similarly. Trial 1 memory for past performance is associated with Trial 2 JOLs	Ariel & Dunlosky, 2011; Hertzog, Hines, & Touron, 2013; Serra & Ariel, 2014	YAs and OAs are expected to spontaneously think about past performance for that item when presented with that item again
Overall task difficulty	Cognitive tasks with higher cognitive load result in more reported TRI	McVay et al., 2013; Zavagnin et al., 2014	Laboratory tasks that have higher cognitive demands should increase spontaneous monitoring

YA younger adults, *OA* older adults, *JOL* judgments of learning

Situational predictors and individual influences of spontaneous monitoring

In this review, we have touched upon different task-related or situational factors that influence metacognitive monitoring accuracy. Individuals attend to and process different types of task-related cues when making predictions about task performance. Manipulating the accessibility of different cues alters monitoring accuracy, but it is unknown whether manipulating the accessibility of different cues will alter how frequently individuals think about their task performance. Examples of common metacognitive cues, along with potential approaches to examine the effect of these cues on monitoring frequency, are shown in Table 1.

Table 2. Potential individual-level influences on spontaneous monitoring

Individual-level	Effect on monitoring accuracy	Relevant citations	Potential effect on monitoring frequency
Age	OAs report more TRI	McVay et al., 2013; Zavagnin et al., 2014; Frank et al., 2015; Jordano & Touron, 2017	At least for more challenging tasks, OAs should have increased spontaneous monitoring relative to YAs
Personality (agreeableness, conscientiousness, openness, extraversion)	YAs higher extraversion, openness, and conscientiousness are overconfident in their predictions. OAs higher in agreeableness and conscientiousness have more accurate judgments	Pallier et al., 2002; Schaefer et al., 2004; Dahl et al., 2010; Schaefer et al., 2004; Stankov & Kleitman, 2008; Burratti, Allwood, & Kleitman, 2013; Hülur, Hertzog, Pearman, & Gerstorf, 2015	Measuring correlations between scores on personality assessments and monitoring frequency may be an area of future study. YAs higher extraversion, openness, and conscientiousness may report less spontaneous monitoring. OAs higher in agreeableness and conscientiousness may report more spontaneous monitoring
Motivation	Higher task-related motivation results in fewer TUTs. Effect of motivation plays on TRI is less clear	Unsworth & McMillan, 2013; Frank et al. 2015	High motivation participants may spontaneously monitor more than low motivation participants
Dispositional mindfulness	Higher dispositional mindfulness may be associated with fewer TUTs and more TRI	Fountain-Zaragoza et al., 2016; Frank et al., 2015	Participants higher in dispositional mindfulness may report more spontaneous monitoring
Self-efficacy and control beliefs	Low self-efficacy is associated with disconnect between subjective and objective memory performance in OAs. Low control beliefs are associated with less TRI in YAs and OAs	Cavanaugh & Poon, 1989; Hultsch, Hertzog, & Dixon, 1987; Hertzog et al., 1990	Individuals high in self-efficacy and control beliefs should engage in more spontaneous monitoring.
Mood and affect	Higher negative mood associated with more TUTs and TRI. Individuals with higher negative affect should spontaneously monitor their performance more	Smallwood et al., 2009	Priming of negative mood should result in increased spontaneous monitoring

TRI task-related interference, *TUT* , *YA* younger adults, *OA* older adults

In addition to reviewing the effects of metacognitive cues on monitoring accuracy, we have also briefly discussed different individual-level or dispositional characteristics that influence monitoring accuracy. Such characteristics include personality type, executive control capabilities, task self-efficacy, and other beliefs about the self and the cognitive task being performed. Again, these characteristics often correlate with monitoring accuracy and differences in these factors across participants contribute to individual differences in monitoring accuracy. Although individual-level influences of monitoring accuracy have received some investigation, it is unknown how these characteristics influence monitoring frequency or one's propensity to think about their task performance. Individual-level and dispositional influences on monitoring accuracy, along with potential approaches to examine the effect of these factors on monitoring frequency, are shown in Table 2.

Conclusions

It is important that we understand how frequently and under what circumstances people engage in metacognitive monitoring, but research that provides insight into these questions is relatively sparse and disconnected. Most studies of metacognitive monitoring focus on monitoring accuracy and factors that influence monitoring accuracy. Common models and theories of metacognition assume that individuals naturally and spontaneously engage in thinking about their task performance when they complete various cognitive tasks, but this assumption has not been well-tested and it is unclear how frequently people engage in thoughts regarding their task performance on both laboratory and everyday cognitive tasks.

Monitoring accuracy is influenced by many aspects of cognitive tasks along with characteristics of the individual completing the cognitive task. Individuals attend to various cues when they think about their task performance. Work examining metacognitive cues assumes that individuals spontaneously attend to various cues but, again, work examining cue-utilization during performance monitoring does so by manipulating cues and examining the effects this has on the accuracy of individuals' explicit metacognitive judgments. Examining how the accessibility of these cues affects the frequency of spontaneous monitoring can provide us with additional information regarding how participants naturally attend to various metacognitive cues. Additionally, individual differences in cognitive capacity, personality, and beliefs about cognition appear to influence monitoring accuracy. We know that monitoring judgments are more accurate for some individuals, but it is unclear whether these same factors also influence (or reflect) the quantity or frequency of individuals' monitoring.

Examining the frequency of spontaneous monitoring can provide critical answers to questions regarding the relationships between performance monitoring, task performance, and metacognitive control. Metacognitive accuracy is linked to better task performance and, if individuals who more frequently engage in monitoring tend to have better monitoring accuracy, then one way to increase task performance in younger and older adults may be to encourage or train individuals to more frequently check in on their task performance.

Focused studies examining spontaneous monitoring may also provide clues as to why older adults show spared monitoring accuracy, despite age-related declines in processing speed and working memory performance. One reason for older adults' spared monitoring may be that they engage in more spontaneous monitoring and thus have more expertise with metacognitive monitoring. This review describes specific directions for future research that will help to address these fundamental questions regarding the frequency and consequences of spontaneous metacognition.

References

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences, 103*(24), 9369-9372. https://doi.org/10.1073/pnas.0601071103

Antrobus, J. S., Singer, J. L., Goldstein, S., & Fortgang, M. (1970). Mind wandering and cognitive structure. *Transactions of the New York Academy of Sciences, 32*(2), 242–252. https://doi.org/10.1111/j.2164-0947.1970.tb02056.x

Azevedo, R., D. C. Moos, A. M. Johnson, & A. D. Chauncey. (2010). "Measuring Cognitive and Metacognitive Regulatory Processes During Hypermedia Learning: Issues and Challenges." *Educational Psychologist, 45*(4), 210–223. https://doi.org/10.1080/00461520.2010.515934

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology, 81*(1), 126-131. https://doi.org/10.1037/h0027455

Ariel, R., & Dunlosky, J. (2011). The sensitivity of judgment-of-learning resolution to past test performance, new learning, and forgetting. *Memory and Cognition, 39*(1), 171-184. https://doi.org/10.3758/s13421-010-0002-y

Bailey, H., Dunlowsky, J., & Hertzog, C. (2010). Metacognitive training at home: Does it improve older adults' learning? *Gerontology, 56*(4), 414–420.

Baker, J., Dunlosky, J., & Hertzog, C. (2010). How accurately can older adults evaluate the quality of their text recall? The effect of providing standards on judgment accuracy. *Applied Cognitive Psychology, 24*(1), 134. https://doi.org/10.1002/acp.1553

Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. *Language Testing*, *28*(1), 51–75. https://doi.org/10.1177/0265532210376379

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28*(5), 610-632. https://doi.org/10.1016/0749596X(89)90016-8

Bottiroli, S., Dunlosky J., Guerini, K., Cavallini, E., & Hertzog C. (2010). Does task affordance moderate age-related deficits in strategy production? *Aging, Neuropsychology, and Cognition, 17,* 591–602.

Brigham, M.C. & Pressley, M. (1988). Cognitive monitoring and strategy choice in younger and older adults. *Psychology and Aging, 3*(3), 349–257. https://doi.org/10.1037/0882-7974.3.3.249

Brunswick, E. (1943). Organismic achievement and environmental probability. *Psychological Review, 50,* 255-272.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124,* 1–38. https://doi.org/10.1196/annals.1440.011

Buratti, S., Allwood, C.M., & Kleitman, S. (2013). First- and second-order metacognitive judgments of semantic memory reports: The influence of personality traits and cognitive styles. *Metacognition Learning, 8*(1), 79-102. https://doi.org/10.1007/s11409-013-9096-5

Butterfield, E. C., Nelson, T. O., & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology, 24*(5), 654–663. https://doi.org/10.1037/0012-1649.24.5.654

Carstensen, L. L. (1995). Evidence for a life-span theory of socioemotional selectivity. *Current Directions in Psychological Science, 4*(5), 51-56. https://doi.org/10.1111/1467-8721.ep11512261

Castel, A. D., McCabe, D. P., & Roediger, H. L. (2007). Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bulletin and Review, 14*(1), 107-111. https://doi.org/10.3758/BF03194036

Cavanaugh, J.C. & Poon, L.W. (1989). Metamemorial predictors of memory performance in young and older adults. *Psychology and Aging, 4*(3), 365–368. https://doi.org/10.1037/0882-7974.4.3.365

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12*(1), 50-71. https://doi.org/10.1037/08827974.12.1.50

Cook, S.E., & Marsiske, M. (2006). Subjective memory beliefs and cognitive performance in normal and mildly impaired older adults. *Aging and Mental Health, 10*(4), 413–423. https://doi.org/10.1080/13607860600638487

Coutinho, M., Redford, J., Church, B., Zakrzewski, A., Couchman, J., & Smith, D. (2015). The interplay between uncertainty monitoring and working memory: Can metacognition become automatic? *Memory and Cognition, 43*(7), 990-1006. https://doi.org/10.3758/s13421-015-0527-1.

Coy, B., O'Brien, W.H., Tabaczynski, T., Northern, J., & Carels, R. (2011). Associations between evaluation anxiety, cognitive interference and performance on working memory tasks. *Applied Cognitive Psychology, 25*(5), 823-832. https://doi.org/10.1002/acp.1765

Craik, F.I.M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society, 302*(1110), 341–359. https://doi.org/10.1098/rstb.1983.0059

Craik, F. I. M., & Byrd M. (1982). Aging and cognitive deficits: The role of attentional resources. In Craik F. I. M. & Trehub S., (Eds.), Aging and cognitive processes (pp. 191–211). New York, NY: Plenum.

Crumley, J.J., Stetler C.A., & Horhota, M. (2014). Examining the relationship between subjective and objective memory performance in older adults: A meta-analysis. *Psychology and Aging, 29*(2), 250–263. https://doi.org/10.1037/a0035908

Cull, W. L., & Zechmeister, E. B. (1994). *Memory & Cognition, 22*(2), 249–257. https://doi.org/10.3758/BF03208896.

Dahl, M., Allwood, C. M., Rennemark, M., & Hagberg, B. (2010). The relation between personality and the realism in confidence judgements in older adults. *European Journal of Ageing, 7*(4), 283-291. https://doi.org/10.1007/s10433-010-0164-2

Damoiseaux, J., Beckmann, C., Arigita, E., Barkhof, F., Scheltens, P., Stam, C., Smith, S., & Rombouts, S. (2008). Reduced resting-state brain activity in the "default network" in normal aging. *Cerebral Cortex, 18*(8), 856–1864. https://doi.org/10.1093/cercor/bhm207

Deekens, V.M., Greene, J.A., & Lobczowski, N.G. (2017). Monitoring and depth of strategy use in computer-based learning environments for science and history. *British Journal of Educational Psychology.* https://doi.org/10.1111/bjep.12174

Deekens, V. M., Greene, J. A., & Lobczowski, N. G. (2018). Monitoring and depth of strategy use in computer based learning environments for science and history. *The British Journal of Educational Psychology, 88*(1), 63–79. https://doi.org/10.1111/bjep.12174.

Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., ... Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly*, *50*(4), 393-416. https://doi.org/10.1002/rrq.105

Dodson, C.S., Bawa, S., Krueger, L.E. (2007). Aging, metamemory, and high-confidence errors: a misrecollection account. *Psychology and Aging, 22*(1), 122–133. https://doi.org/10.1037/0882-7974.22.1.122

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology*, 47, 274-296.

Dunlosky, J., & Connor, L.T. (1997). Age differences in the allocation of study time account for age differences in memory performance. *Memory and Cognition, 25*(5), 691-700. https://doi.org/10.3758/BF03211311

Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging, 15*(3), 462–474. https://doi.org/10.1037//0882-7974.15J.4

Dunlosky, J., & Metcalfe, J. (2009). Metacognition. San Francisco, CA: Sage Publications.

Eakin, D.K., Hertzog, C., & Harris, W. (2014). Age invariance in semantic and episodic metamemory: Both younger and older adults provide accurate feeling-of-knowing for names of faces. *Aging, Neuropsychology, and Cognition, 21*(1), 27–51. https://doi.org/10.1080/13825585.2013.775217

Ericsson, K. A. & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: The MIT Press.

Ericsson, K.A. & Simon, H., (1987). Verbal reports on thinking. In: Færch, C., Kasper, G. (Eds.), Introspection in Second Language Research. Multilingual Matters, Clevedon, pp. 24-53.

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

Fernandez J, & Jamet E. (2017). Extending the testing effect to self regulated learning. *Metacognition and Learning, 12,* 131-156.

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source relevant cues. *Psychology and Aging, 7*(3), 443-452. https://doi.org/10.1037//0882-7974.7.3.443

Fidler, E. L. (1983). The reliability and validity of concurrent,retrospective, and interpretive verbal reports: An experimental study. In P. Humphreys, O. Svenson, & A. Vari (Eds.), Analyzing and aiding decision processes (pp. 429-440). Amsterdam: North-Holland

Flavell, J.H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34*(10), 906-911. https://doi.org/10.1037/003-066X.34.10.906

Forster, S. & Lavie N. (2008). Attentional capture by entirely irrelevant distractors. *Visual Cognition, 16*(2), 200–214. https://doi.org/10.1080/13506280701465049

Fountain-Zaragoza, S., Londerée, A., Whitmoyer, P., & Prakash, R. S. (2016). Dispositional mindfulness and the wandering mind: implications for attentional control in older adults. *Consciousness and Cognition, 44,* 193–204. https://doi.org/10.1016/j.concog.2016.08.003

Fox, M. C., Baldock, Z., Freeman, S. P., & Berry, J. M. (2016). The role of encoding strategy in younger and older adult associative recognition: A think-aloud analysis. *Psychology and Aging, 31*(5), 471-480. https://doi.org/10.1037/pag0000111

Frank, D.J., Nara, B., Zavagnin, M., Touron, D.R., & Kane, M.J. (2015). Validating older adults' reports of less mind-wandering: An examination of eye movements and dispositional influences. *Psychology and Aging, 30*(2), 266-278. https://doi.org/10.1037/pag0000031

Fryer, D. H. (1941). Articulation in automatic mental work. *American Journal of Psychology, 54,* 504-517.

Gardner, R. S., & Ascoli, G. A. (2015). The natural frequency of human prospective memory increases with age. *Psychology and Aging, 30*(2), 209–219. https://doi.org/10.1037/a0038876

Garner, R., & Taylor, N. (1982). Monitoring of understanding: An investigation of attentional assistance needs at different grade and reading proficiency level. *Reading Psychology, 1,* 1-6.

Giambra, L. M. (1989). Task-unrelated thought frequency as a function of age: A laboratory study. *Psychology and Aging, 4*(2), 136-143. https://doi.org/10.1037/0882-7974.4.2.136

Giambra, L.M. (1995). A laboratory method for investigating influence on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition, 4,* 1-21. https://doi.org/10.1006/ccog.1995.1001

Gilhooly, K.J., Wynn, V., Phillips, L.H., Logie, R.H., & Della Sala, S. (1999). Visuo-spatial and verbal working memory in the five-disk Tower of London task: an individual differences approach. *Thinking and Reasoning*.

Glaser, R. (1991). Expertise and assessment. In Wittrock, M., Baker, E. (Eds.), Testing and cognition (pp. 17–30). Englewood Cliffs, NJ: Prentice-Hall.

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.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56*(4), 208–216. https://doi.org/10.1037/h0022263

Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior, 6*(5), 685–691. https://doi.org/10.1016/S0022-5371(67)80072-0

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), Attention and performance XVII (pp. 653–675). Cambridge, MA: MIT Press.

Hertzog, C., & Hultsch, D. F. (2000). Metacognition in adulthood and old age. In F. I. M. Craik & T. A. Salthouse (Eds.), The handbook of aging and cognition (2nd ed., pp. 417–466). Mahwah, NJ: Lawrence Erlbaum.

Hertzog, C., & Dunlosky, J. (2011). Metacognition in later adulthood: Spared monitoring can benefit older adults' self-regulation. *Current Directions in Psychological Science, 20*(3), 167-173. https://doi.org/10.1177/0963721411409026

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging,5*(2), 215-227.

Hertzog, C., Dunlosky, J., Robinson, E. A., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 22–34. https://doi.org/10.1037/0278-7393.29.1.22.

Hertzog, C., Hines, J. C., & Touron, D. R. (2013). Judgments of learning are influenced by multiple cues in addition to memory for past test accuracy. *Archives of Scientific Psychology*, *1*(1), 23-32. https://doi.org/10.1037/arc0000003

Hertzog, C., Price, J., & Dunlosky, J. (2008). How is knowledge generate about memory encoding strategy effectiveness. *Learning and Individual Differences, 18*, 430–445. https://doi.org/10.1016/j.lindif.2007.12.002.

Hertzog, C., Sinclair, S. M., & Dunlosky, J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. *Developmental Psychology, 46*(4), 939–948. https://doi.org/10.1037/a0019812.

Hertzog, C., Touron, D.R., & Hines, J. (2007). Does a time monitoring deficit influence older adults' delayed retrieval shift during skill acquisition? *Psychology and Aging, 22*(3), 607-624. https://doi.org/10.1037/0882-7974.22.3.607

Hertzog, C., Saylor, L. L., Fleece, A. M., & Dixon, R. A. (1994). Metamemory and aging: Relations between predicted, actual and perceived memory task performance. *Aging and Cognition, 1*(3), 203–237. https://doi.org/10.1080/13825589408256577

Hertzog, C., Kidder, D.P., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? *Psychology and Aging, 17*(2), 209–225. https://doi.org/10.1037/0882-7974.17.2.209

Hultsch, D.F., Hertzog, C., & Dixon, R.A. (1987). Age differences in metamemory: resolving the inconsistencies. *Canadian Journal of Psychology, 41*(2), 193–208. https://doi.org/10.1037/h0084153

Hülür, G., Hertzog, C., Pearman, A.M., & Gerstorf, D. (2015). Correlates and moderators of change in subjective memory and memory performance: Findings from the Health and Retirement Study. *Gerontology, 61*(3), 232-240. https://doi.org/10.1159/000369010

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students: A think aloud study. *European Journal of Psychology of Education*, *21*, 35–52.CrossRefGoogle Scholar

Jordano, M. J., & Touron, D. R. (2017a). Stereotype Threat as a Trigger of Mind-Wandering in Older Adults. *Psychology and Aging.* https://doi.org/10.1037/pag0000167.

Jordano, M.J., & Touron, D.R. (2017b). Priming performance-related concerns induces task related mind-wandering. *Consciousness and Cognition*, 55, 126-135. https://doi.org/10.1016/j.concog.2017.08.002.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18*(7), 614-621. https://doi.org/10.1111/j.1467-9280.2007.01948.x

Karpf, D. A. (1973). Thinking aloud in human description learning. *Dissertation Abstracts International, 33,* 6111-B. (University Microfilms No. 73-13625.)

Kelemen, W.L., Frost, P.J., & Weaver III, C.A. (2000). Individual differences in metacognition: evidence against a general metacognitive ability. *Memory and Cognition, 28*(1), 92–107. https://doi.org/10.3758/BF03211579

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*(4), 704–721. https://doi.org/10.1016/S0749596X(02)00504-1

Kempe, M., Kalicinski, M., & Memmert, D. (2015). Naturalistic assessment of everyday memory performance among older adults. *Experimental Aging Research*, *41*(4), 426-445. https://doi.org/10.1080/0361073X.2015.1053766

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330,* 932.

Klinger, E. (1971). Structure and functions of fantasy. New York: Wiley.

Klinger, E. (1978). “Modes of normal conscious flow”. In: K.S. Pope & J.L. Singer (Eds). The stream of consciousness: Scientific investigations into the flow of human experience. New York: Plenum.

Klinger, E., & Cox, W. M. (1987-88). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality, 7*(2), 105-128. https://doi.org/10.2190/7K24-G343-MTQW-115V

Komori, M. (2016). Effects of working memory capacity on metacognitive monitoring: A study of group differences using a listening span test. *Frontiers in Psychology, 7,* 285. https://doi.org/10.3389/fpsyg.2016.00285

Kopp, K., D'Mello, S., & Mills, C. (2015). Influencing the occurrence of mind wandering while reading. *Consciousness and Cognition, 34*(1), 52–62. https://doi.org/10.1016/j.concog.2015.03.003

Koriat, A. (1997). Monitoring one's knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Lachman, M.E., & Agrigoroaei, S. (2012). Low perceived control as a risk factor for episodic memory: The mediational role of anxiety and task interference. *Memory and Cognition, 40*(2), 287–296. https://doi.org/10.3758/s13421-011-0140-x

Lachman, M.E., Neupert, S.D., Agrigoroaei, S. (2011). The relevance of control beliefs for health and aging. In: K.W. Schaie & S.L. Willis (Eds.). Handbook of the psychology of aging (pp. 175-190). Burlington, MA: Academic Press.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 464-470. https://doi.org/10.1037/0278-7393.16.3.464

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition, 26,* 35-57. https://doi.org/10.1017/S0272263104026129.

Lovelace, E. A., & Marsh, G. R. (1985). Prediction and evaluation of memory performance by young and old adults. *Journal of Gerontology, 40*(2), 192–197. https://doi.org/10.1093/geronj/40.2.192

Lumley, T. (2005). Assessing second language writing: The rater's perspective. Peter Lang: Frankfurt.

Maillet, D., & Rajah, M. N. (2013). Age-related changes in frequency of mind-wandering and task-related interferences during memory encoding and their impact on retrieval. *Memory, 21*(7), 818–831. https://doi.org/10.1080/09658211.2012.761714.

Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 663-679. https://doi.org/10.1037//0278-7393.10.4.663

Martin, S. (1987). The meaning-making strategies reported by proficient readers and writers. Paper presented at the meeting of the National Reading Conference St. Petersburg, FL.

Masicampo, E. J., & Baumeister R. F. (2011). Consider it done! Plan making can eliminate the cognitive effects of unfulfilled goals. *Journal of Personality and Social Psychology, 101*(4), 667-683. https://doi.org/10.1037/a0024192

Matthews, G., Joyner, L., Gilliland, K., Huggins, J., & Falconer, S. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three? In I. Merville, I. J. Deary, F. DeFruyt, & F. Ostendorf (Eds.), Personality psychology in Europe (pp. 335-350. Tilburg: Tilburg University Press.

McDonald-Miszczak, L., Hunter, M. A., & Hultsch, D. F. (1994). Adult age differences in predicting memory performance: The effects of normative information and task experience. *Canadian Journal of Experimental Psychology, 48*(1), 95-118.

McDowd, J. M., & Craik, F. I. M. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance, 14*(2), 267–280. https://doi.org/10.1037/0096-1523.14.2.267

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind-wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 196–204. https://doi.org/10.1037/a0014104.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin, 136*(2), 188–197. https://doi.org/10.1037/a0018298

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "D'oh!": working memory capacity and mind wander-ing predict extreme reaction times and executive control errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(3), 525-549. https://doi.org/10.1037/a0025896

McVay, J. C., & Kane, M. J. (2013). Dispatching the wandering mind? Toward a laboratory method for cuing "spontaneous" off-task thought. *Frontiers in Psychology, 4*, 570. https://doi.org/10.3389/fpsyg.2013.00570.

McVay, J. C., Meier, M. E., Touron, D. R., & Kane, M. J. (2013). Aging ebbs the flow of thought: Adult age differences in mind wandering, executive control, and self-evaluation. *Acta Psychologica, 142*(1), 136–147. https://doi.org/10.1016/j.actpsy.2012.11.006

Meyers, J., Lytle, S., Palladino, D., Devenpeck, G. & Green, M. (1990). Think-aloud protocol analysis: An investigation of reading comprehension strategies in fourth- and fifth-grade students. *Journal of Psychoeducational Assessment, 8,* 112–127.

Mokos E, & Kafoussi S. (2013). Elementary students' spontaneous metacognitive functions in different types of mathematical problems. *Journal of Research in Mathematics Education, 2*(2): 242–267. https://doi.org/10.4471/redimat.2013.29.

Montague, M., Applegate, B. (1993). Middle school students' mathematical problem solving: An analysis of think-aloud protocols. *Learning Disability Quarterly, 16*(1), 19–31. https://doi.org/10.2307/1511157

Mrazek, M.D., Chin, J.M., Schmader, T., Hartson, K.A., Smallwood, J., & Schooler, J.W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology, 47*(6), 1243-1248. https://doi.org/10.1016/j.jesp.2011.05.011

Murphy, M.D., Sanders, R.E., Gabriesheski, A.S., & Schmitt, F.A. (1981). Metamemory in the aged. *Journal of Gerontology, 36*(2)*,* 185-193.

Murphy, M.D., Schmitt, F.A., Caruso, M.J., Sanders, R.E. (1987). Metamemory in older adults: The role of monitoring in serial recall. *Psychology and Aging, 2,* 331–339.

Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology, 91,* 488–496.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*(2), 102-116. https://doi.org/10.1037/0003-066X.51.2.102

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*(4), 676-686. https://doi.org/10.1037/0278-7393.14.4.676

Nelson, T. O., Leonesio, R. J., Shimamura, A. P., Landwehr, R. S., & Narens, L. (1982). Overlearning and the feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(4), 279-288. https://doi.org/10.1037/0278-7393.8.4.279

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. 125–173). New York: Academic Press.

Owings, R., Petersen, G., Bransford, D. D., Morris, C. D., & Stein, B. S. (1980). Spontaneous monitoring and regulation of learing: A comparison of successful and less successful fifth-graders. *Journal of Educational Psychology, 72.* 250–256.

Pallier, G., Wilkinson, R., Danthir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology, 129*(3), 257-299. https://doi.org/10.1080/00221300209602099

Park, D. C., Hertzog, C., Leventhal, H., Morrell, R. W., Leventhal, E., Birchmore, D., Martin, M., & Bennett, J. (1999). Medication adherence in Rheumatoid arthritis patients: Older is wiser. *Journal of the American Geriatrics Society*, *47*(2), 172–183. https://doi.org/10.1111/j.1532-5415.1999.tb04575.x

Peterson, P. L., Swing, S. R., Braverman, M. T., & Buss, R. R. (1982). Students' aptitudes and their reports of cognitive processes during direct instruction. *Journal of Educational Psychology*, *74*(4), 535-547. https://doi.org/10.1037/0333-0663.74.4.535

Rabinowitz, J.C., Ackerman, B.P., Craik, F.I.M., & Hinchley, J.L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology, 37*(6), 688-695. https://doi.org/10.1093/geronj/37.6.688

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*(1), 90–138. https://doi.org/10.1016/0010-0285(87)90005-3.

Rendell, P. G., & Thomson, D. M. (1999). Aging and prospective memory: Differences between naturalistic and laboratory tasks. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *54*(4), 256-269. https://doi.org/10.1093/geronb/54B.4.P256

Robinson, A.E., Hertzog C., & Dunlosky J. (2006). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, and Cognition, 13*(3-4), 458–478. https://doi.org/10.1080/13825580600572983

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols, *Memory and Cognition, 17 (November),* 759-769.

Sarason, I.G., Sarason, B.R., Keefe, D.E., Hayes, B.E. & Shearin, E.N. (1986) Cognitive interference: Situational determinants and trait-like characteristics. *Journal of Personality and Social Psychology, 51*(1), 215-226. https://doi.org/10.1037/0022-3514.51.1.215

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2*(3), 179-183.

Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, W. K. (2004). Overconfidence and the Big Five. *Journal of Research in Personality, 38*(5), 473-480. https://doi.org/10.1016/j.jrp.2003.09.010

Schellings, G. L. M., & Broekkamp, H. (2011). Signaling task awareness in think-aloud protocols from students selecting relevant information from text. *Metacognition Learning, 6,* 65–82. https://doi.org/10.1007/s11409-010-9067-z

Schellings, G.L. M., van Hout-Wolters, B.A. H. M., Veenman, M.V. J., & Meijer, J. (2013). Assessing metacognitive activities: the in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European Journal of Psychology of Education,*1–28. https://doi.org/10.1007/s10212-012-0149-y.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 336–356. https://doi.org/10.1037/0033-295X.115.2.336

Schoenfeld, A. H. (1985). Mathematical problem solving. Orlando, FL: Academic Press.

Schoenfeld, A.H. (1987) What's All the Fuss about Metacognition? In A.H. Schoenfeld (Ed.) Cognitive science and mathematics education, (pp. 189-215). Hillsdale: Lawrence Erlbaum.

Seli, P., Carriere, J.S.A., Levene, M., & Smilek, D. (2013). How few and far between? Examining the effects of probe rate on self-reported mind wandering. *Frontiers in Psychology, 4,* 430. https://doi.org/10.3389/fpsyg.2013.00430

Serra, M. J., & Ariel, R. (2014). People use the memory for past test heuristic as an explicit cue for judgments-of-learning. *Memory and Cognition, 42*(8), 1260-1272. https://doi.org/10.3758/s13421-014-0431-0.

Shanks**,** L. L., & Serra, M. J. (2014). Domain familiarity as a cue for judgments of learning. *Psychonomic Bulletin and Review, 21*(2), 445-453. https://doi.org/10.3758/s13423-013-0513-1

Shaw, R.J., & Craik, F.I.M. (1989). Age differences in predictions and performance on a cued recall task. *Psychology and Aging, 4*(2), 131–135.

Simmons, J. P., & Nelson, L. D. (2006). Intuitive confidence: Choosing between intuitive and nonintuitive alternatives. *Journal of Experimental Psychology: General, 135*(3), 409-428. https://doi.org/10.1037/0096-3445.135.3.409

Smagorinsky, P. (1994) Think-aloud protocol analysis: Beyond the black box. In P. Smagorinsky (ed.), Speaking about writing: Reflections on research methodology (pp. 3–19). Thousand Oaks, CA: Sage.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958. https://doi.org/10.1037/0033-2909.132.6.946

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: negative moods lead the mind to wander. *Emotion, 9*(2), 271–276. https://doi.org/10.1037/a0014855

Smallwood, J., O'Connor, R. C., & Heim, D. (2005). Rumination, dysphoria, and subjective experience. *Imagination, Cognition, and Personality, 24*(4), 355–367. https://doi.org/10.2190/AE18-AD1V-YF7L-EKBX

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13(*4), 657-690. https://doi.org/10.1016/j.concog.2004.06.003

Souchay, C., Isingrini, M., Clarys, D., Taconnat, L., & Eustache, F. (2004). Executive functioning and judgment-oflearning versus feeling-of knowing in older adults. *Experimental Aging Research, 30*(1), 47–62. https://doi.org/10.1080/03610730490251478.

Stankov, L., & Kleitman, S. (2008). Processes on the borderline between cognitive abilities and personality: Confidence and its realism. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE hand-book of personality theory and assessment* (Vol. 1, pp. 545–559). Los Angeles: SAGE.

Stawarczyk, D., Majerus, S., Maj, M., Van Der Linden, M., & D'Argembeau, A. (2011a). Mind wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica, 136*(3), 370–381 https://doi.org/10.1016/j.actpsy.2011.01.002

Stawarczyk, D., Majerus, S., Maquet, P., & D'Argembeau, A. (2011b). Neural correlates of ongoing conscious experience: Both task-unrelatedness and stimulus-independence are related to default network activity. *PLoS ONE, 6*(2), e16997. https://doi.org/10.1371/journal.pone.0016997

Steele, C. M., & Aronson, J. (1995). Contending with a stereotype: African-American intellectual test performance and stereotype threat. *Journal of Personality and Social Psychology, 69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Stine-Morrow, E. A. L., Shake, M. C., Miles, J. R., & Noh, S. R. (2006). Adult age differences in the effects of goals on self-regulated sentence processing. *Psychology and Aging, 21*, 790-803.

Teasdale, J.D., Dritschel, B.H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory and Cognition, 23*(5), 551–559. https://doi.org/10.3758/BF03197257

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes, 21*(3), 255–288. https://doi.org/10.1080/01638539609544959.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Unsworth, N., Redick, T.S., Lakey, C.E., & Young, D.L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence, 38*(1), 111–122. https://doi.org/10.1016/j.intell.2009.08.002

Undorf, M., & Erdfelder, E. (2013). Separation of encoding fluency and item difficulty effects on judgements of learning. *The Quarterly Journal of Experimental Psychology, 66*(10), 2060-2072. https://doi.org/10.1080/17470218.2013.777751

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(3), 832–842. https://doi.org/10.1037/a0029669.

Veenman, M.V.J., Elshout, J.J., & Groen, M.G.M. (1993). Thinking aloud: Does it affect regulatory processes in learning?. *Tijdschrift voor Onderwijsresearch, 18*, 322-330.

Verhaeghen, P., Martin, M., & Sędek, G. (2012). Reconnecting cognition in the lab and cognition in real life: The role of compensatory social and motivational factors in explaining how cognition ages in the wild. *Neuropsychology, Development, and Cognition: Section B, Aging, Neuropsychology and Cognition*, *19*(1-2), 1–12. https://doi.org/10.1080/13825585.2011.645009

Wahlin, A., Maitland, S. B., Bäckman, L., & Dixon, R. A. (2003). Interrelationships between subjective health and episodic memory change in Swedish and Canadian samples of older adults. *International Journal of Aging and Human Development, 57*(1), 21–35. https://doi.org/10.2190/9VAA-KMYV-U2HU-PVAW

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

Yuill, N. & Oakhill, J. (1991). Children's problems in text comprehension. Cambridge: Cambridge University Press.

Zelinski, E. M., Gilewski, M. J., & Thompson, L. W. (1980). Do laboratory tests relate to self assessment of memory ability in the young and old? In L.W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.). New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference (pp. 519-538). Hillsdale, NJ: Lawrence Erlbaum.

Zavagnin, M., Borella, E., & De Beni, R. (2014). When the mind wanders: Age-related differences between young and older adults. *Acta Psychologica, 145*(1), 54–64. https://doi.org/10.1016/j.actpsy.2013.10.016