JORDANO, MEGAN J., Ph.D. How Often Do Younger and Older Adults Engage in Monitoring? A New Approach to Studying Metacognition. (2018)
Directed by Dr. Dayna R. Touron. 134 pp.

Older adults demonstrate spared metacognitive monitoring abilities, despite cognitive decline in other domains. An extensive literature examines *how accurately* individuals engage in monitoring. The question of *how often* individuals engage in metacognition has been neglected. It is also possible that individuals who monitor more often also monitor more accurately, and age-related *increases* in monitoring contributes to older adults' intact monitoring abilities. In the current study, younger and older adults were assigned to one of two conditions. Control condition participants completed a learning task containing thought content probes. Experimental condition participants completed the same learning task containing both thought content probes and screens asking them to make *judgments of learning* (JOLs). This design allows us to compare monitoring frequency in younger and older adults, determine how making explicit metacognitive judgments alters propensity to engage in monitoring, and examine the relationship between monitoring frequency and monitoring accuracy within the experimental condition. Older adults engaged in more frequent monitoring than younger adults. Additionally, older adults who were required to provide JOLs engaged in more frequent monitoring than older adults who were not required to make JOLs. Finally, younger and older adults who engaged in more frequent monitoring were not found to have more accurate metacognitive judgments than those who engaged in less frequent monitoring.

HOW OFTEN DO YOUNGER AND OLDER ADULTS ENGAGE IN MONITORING?

A NEW APPROACH TO STUDYING METACOGNITION

by

Megan J. Jordano

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2018

Approved by

Dayna R. Touron
Committee Chair

APPROVAL PAGE

This dissertation, written by Megan J. Jordano has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair Dr. Dayna R. Touron

Committee Members Dr. Michael J. Kane

Dr. Ethan Zell

Dr. Christopher Wahlheim

03/20/18
Date of Acceptance by Committee

03/5/18
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Imagine you are attending a lecture. It is important to you that you pay close attention to the information the speaker is conveying and that you can later recall this information. As you listen, you find yourself engaging in extraneous thoughts. You might realize you are getting distracted by noises in your environment and that you need to refocus your attention. Or perhaps you periodically summarize or repeat to yourself what the speaker had just said. If you find your comprehension lacking, you may determine that you should ask the speaker to repeat or explain a critical piece of information. These are just some examples of *metacognition*, or thoughts one has regarding their own cognitive processes (Flavell, 1979).

Three distinct yet interrelated 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 neglected within the study of metacognition.

An extensive literature has examined metacognitive monitoring, generally focusing on whether individuals' judgments about their task performance *accurately* reflect how well they are performing a task of interest, and which factors predict or influence the accuracy of one's performance-related thoughts. For instance, a PsychInfo search including the terms "metacognition" and "monitoring" results in well over a thousand separate sources. The question of *how frequently* individuals spontaneously monitor their task performance is an also critical but thus far 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 of particular interest. Older adults' metacognitive monitoring abilities remain relatively spared despite declines in other types of cognitive functioning (Shaw & Craik, 1989; Salthouse, 1991; Hasher, Zacks, & May, 1999; Dunlosky & Connor, 1997) but no work within the metacognition literature has examined the specific *frequency* and *content* of monitoring experiences in younger and 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*. Following this review of the monitoring literature, a potential method from the mind-wandering literature that can be used to measure the frequency of spontaneous metacognitive thoughts is discussed, as well as findings from the mind-wandering literature that may provide an indication of how frequently individuals think about their task performance. Finally, we describe and discuss the results of a study designed to examine the frequency of spontaneous metacognitive monitoring in younger and older adults, the effect of making metacognitive judgements on the frequency and content of mind-wandering, and the relationship between spontaneous metacognitive monitoring and metacognitive monitoring accuracy. This study hopefully provides a starting point for future studies examining monitoring frequency and factors that influence how often individuals engage in monitoring.

Measuring Monitoring Accuracy

The relationships between cognition and different types of metacognitive thoughts can be understood by examining popular models of metacognition, such as the Nelson and Narens model (1990) 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 specifically ask them to reflect on their task performance.

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 performance improved. 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 after one has acquired new information but before that information has been tested. JOLs 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 during test and reflect participants' thoughts about how well they were able to recall information from memory.

In metacognition studies, participants are often asked to make their metacognitive judgments during learning and memory tasks such as the paired associates task (Arbuckle & Cuddy, 1969; Leonesio & Nelson, 1990; Nelson & Dunlosky, 1991; Dunlosky & Nelson, 1994), in which participants study pairs of words in anticipation of a later memory test. The accuracy of performance monitoring can be measured in terms of relative accuracy or absolute accuracy. *Relative accuracy* is typically measured using discrimination indices such as gamma correlations (Goodman & Kruskal, 1954; Nelson, 1984), which assess the degree to which a person's performance judgments distinguish between *correctly versus incorrectly* recalled items. *Absolute accuracy* is the degree to which one's level of judgments matches one's actual level of performance (Maki, Shields, Wheeler, & Zacchilli, 2005), and, unlike gamma correlations, typically account for judgment bias. Measures include difference scores and calibration curves.

Monitoring Accuracy and Aging

Normal cognitive aging is characterized by declines in a wide variety of cognitive tasks. Older adults often have poorer performance on memory tasks compared to younger adults (Lovelace & Marsh, 1985; Rabinowitz, Ackerman, Craik, & Hinchley, 1982; Shaw & Craik, 1989) and also report more memory-related complaints than do younger adults (Hertzog & Hultsch, 2000; Levy-Cushman & Abeles, 1998). Cognitive aging is also characterized by 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 (Connor, Dunlosky, & Hertzog, 1997; Halamish, McGillivray, & Castel, 2011; Hertzog, Dixon, & Hultsch, 1990; Hertzog, Kidder, Powell, Moman, & Dunlosky, 2002; Hertzog, Saylor, Fleece, & Dixon, 1994; Robinson, Hertzog, & Dunlosky, 2006; Lovelace & Marsh, 1985; Shaw & Craik, 1989). Absolute JOL accuracy may also be equivalent in younger and older adults (McDonald-Miszczak, Hunter, & Hultsch, 1994).

While there is evidence of age equivalence in monitoring, studies have found age-related deficits in related abilities (Bruce, Coyne, & Botwinick, 1982; Devolder, Brigham, & Pressley, 1990; Lovelace, 1990). Older adults struggle with aggregating metacognitive information, resulting in monitoring deficits when making global predictions of performance (Brigham & Pressley, 1988; Hertzog et al., 2008; Hertzog, Touron, & Hines, 2007). However, older adults are typically as accurate as younger adults when asked to make item-by-item JOLs (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). As with JOLs, younger adults' FOKs are moderately correlated with task performance (Hart, 1967; Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982). Absolute accuracy for FOK judgments about general knowledge also shows age invariance (Butterfield, Nelson, & Peck, 1988; Eakin, Hertzog, & Harris, 2014; Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007). Even in tasks where older adults have worse memory performance, they have FOK resolution similar to younger adults' (Hertzog & Dunlosky, 2011; Hertzog, Dunlosky, & Sinclair, 2010).

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). Additionally, older adults sometimes 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).

Metacognitive Control Across Adulthood

The accuracy of individuals' monitoring processes is a topic of interest in part because accurate evaluations regarding task performance often result in more effective cognitive control. Metacognitive judgements such as JOLs correlate with allocation of self-based study time (Nelson & Leonesio, 1988; Connor, Dunlosky, & Hertzog, 1997), other strategic choices (for a review, see Dunlosky, Hertzog, Kennedy, & Thiede, 2005), and future recall performance (Arbuckle & Cuddy, 1969). Although younger and older adults seem to think about their task performance in similar ways, there is evidence that, despite age-invariance in metacognitive monitoring, younger and older adults do not implement metacognitive control in similar ways (Pansky, Goldsmith, Koriat, & Pearlman-Avnion, 2009; Souchay & Isingrini, 2012; Kuhlmann & Touron, 2011). Older adults report using fewer encoding strategies than younger adults (Loewen, Shaw, & Craik, 1990) and are also less effective than younger adults at monitoring the effectiveness of different learning strategies (Brigham & Pressley, 1988; Pressley, Levin, & Ghatala, 1984), even though they do not seem to have issues in monitoring the products of their memory processes. Age differences in strategy choice and strategy monitoring may contribute to age differences in later memory performance, even if performance monitoring itself is spared.

Mind-Wandering and Spontaneous Metacognitive Monitoring

Particularly relevant to the question of how often individuals spontaneous monitor their task performance is the literature examining intrusive, off-task thoughts that are task-related. Because TRI encompasses thoughts regarding task performance, TRI may be analogous to metacognitive monitoring. Given this, patterns of TRI may provide insight regarding how frequently individuals spontaneously monitor their task performance. Numerous factors have been found to influence the amount of TRI (metacognitive monitoring; McVay et al., 2013; Smallwood et al., 2009; Frank et al., 2015; Jordano & Touron, 2017a), that is experienced during a cognitive task. For example, task difficulty is associated with propensity to engage in mind-wandering, with younger (McVay et al., 2013) and older (Zavagnin et al., 2014) adults reporting more TRI on difficult tasks relative to easier ones. The results of these studies both suggest individuals engage in more metacognitive monitoring when the ongoing task is difficult.

Importantly, age has been found to influence both the amount of probe-caught mind-wandering and the content of mind-wandering experiences. It has been found across various types of cognitive tasks and using both retrospective thought content questionnaires and thought-probes that older adults report less overall mind-wandering than do younger adults (Giambra, 1989; Jackson & Balota, 2012; Krawietz, Tamplin, & Radvansky, 2012; McVay, Meier, Touron, & Kane, 2013; Frank et al., 2015; Jordano & Touron, 2017a), despite age-related declines in working memory capacity and inhibitory processes (Hasher & Zacks, 1988). While they report fewer off-task thoughts about things completely unrelated to the task (task-unrelated thoughts; TUTs) compared to younger adults, older adults do report more TRI than younger adults do (McVay et al., 2013; Zavagnin, Borella, & De Beni, 2014; Frank et al., 2015; Jordano & Touron, 2017a). McVay and colleagues (2013) found that TRI accounted for some of the age-related differences in overall reported mind-wandering during a sustained attention task, but older adults still reported less overall mind-wandering than younger adults. In other words, although older adults report fewer off-task thoughts than younger adults, they have been found to report more thinking about task performance and thus more metacognitive monitoring. Numerous other studies using a variety of different cognitive tasks demonstrate that older adults report less overall mind-wandering than do younger adults (Giambra, 1989; Jackson & Balota, 2012; Krawietz, Tamplin, & Radvansky, 2012; Frank et al., 2015; Jordano & Touron, 2017a). Factors such as more performance-related current concerns being primed by the testing environment (Mrazek et al., 2011; Jordano & Touron, 2017a; Jordano & Touron, 2017b) and one's beliefs regarding their ability to control their cognitive functions (Lachman & Agrigoroaei, 2012) have been found to influence TRI, and may explain older adults' increased TRI relative to younger adults in laboratory settings.

While some studies include TRI as a response option, the exact content of individuals' TRI experiences remains unknown. While we believe that TRI includes performance monitoring (metacognitive monitoring) it is also possible that individuals also think about implementing different task strategies to improve performance (metacognitive control). More work should be done examining the specific content of TRI experiences and the frequency in which individuals spontaneously engage in different types of TRI. Additionally, more work should be conducted to examine how to the frequency and specific content of younger and older adults' TRI experiences differ according to the type of cognitive task being completed.

Although both TUTs and TRI are associated with in-the-moment performance errors (McVay & Kane, 2009; McVay et al., 2013), it is possible that engaging in TRI can benefit task performance. Again, task-related interference may be analogous to metacognitive monitoring, and successfully monitoring performance can lead to better metacognitive control and improved task performance. Whereas some studies have found that TRI is associated with worse task performance (Coy, O'Brien, Tabaczynski, Northern, & Carels, 2011; Smallwood et al., 2004; Smallwood et al., 2006; Smallwood et al., 2009; Jordano & Touron, 2017a), other studies have found that TRI is associated with better task performance (McVay & Kane, 2012) or is not associated with task performance (Stawarczyk, Majerus, Maj, van der Linden, & D'Argembeau, 2011). The relationship between TRI and task performance is normally studied by examining the effect that TRI has on in-the-moment performance or performance for shorter tasks. While TRI may disrupt in-the-moment performance, it is unclear if there are downstream benefits to engaging in TRI that are observable only when using more complex tasks with more potential to strategize. Additional work can be conducted to examine if and when TRI can be beneficial to task performance.

Other Types of Spontaneous Thought

In addition to mind-wandering, other types of spontaneous cognition have been investigated, including involuntary autobiographical memories (Bernsten, 2009), "earworms" or spontaneous reoccurring tunes (Beaman & Williams, 2010), prospective memory (McDaniel & Einstein, 2000), and involuntary semantic memories (Kvavilashvili & Mandler, 2004). Interest in spontaneous cognition has increased in recent years (Andrews-Hanna et al., 2010, Christoff et al., 2011; O'Callaghan, Shine, Lewis, Andrews-Hanna, & Irish, 2015; Smallwood & Schooler, 2015) and the growing mind-wandering literature has led to increased interest in spontaneous cognition within older adults.

Most cognitive aging research has focused on deliberately engaged cognition. However, thoughts may arise spontaneously in the absence of specific task demands or instructions meant to encourage participants to engage in those thought processes. Although there is little research examining spontaneous metacognitive thoughts across the adult life span, there are some studies examining involuntary autobiographical memories, intrusive thoughts, and spontaneous prospective memory. This extant literature on aging and spontaneous cognition can provide some insight into whether one might expect age-related changes in spontaneous thoughts regarding task performance.

As a whole, work on mind-wandering, involuntary autobiographical memories (Berntsen & Rubin, 2002; Moulin et al., 2014; Rubin & Berntsen, 2009), intrusive thoughts (Beadel et al., 2013; Lambert et al., 2013; Magee & Teachman, 2012), and prospective memory within the lab and naturalistic settings (Kvavilashvili & Fisher, 2007; Mullet et al., 2013; Scullin et al., 2011; but see Crovitz & Daniel, 1984) suggest that healthy older adults experience *fewer* spontaneous thoughts than do younger adults. These findings are more consistent with the theories of aging that suggest that older adults have a reduction in cognitive resources that result in them engaging is less self-initiated thinking (Craik, 1983; Craik & Byrd, 1982). Given these findings, it is possible that older adults experience less spontaneous metacognition. However, no work within the metacognition literature has specifically examined spontaneous metacognitive thoughts regarding task performance, and the findings demonstrating increased TRI in older adults instead suggest that spontaneous performance monitoring increases with age.

Measuring Monitoring Frequency

Studying the accuracy of metacognitive thoughts requires participants to make explicit predictions or judgments about their cognitive abilities. While it is likely that participants think about their performance spontaneously and without being prompted to by the experimenter or task instructions, it is also possible that people taking part in metacognition studies engage in monitoring to a lesser or greater degree because they are *explicitly* being asked by the experimenters to make performance judgments. There has been little consensus in regard to the direction of this effect with some studies finding that requiring participants to make metacognitive judgements results in increased task performance and other studies finding that requiring participants to make metacognitive judgements results in decreased task performance (see Double, Birney, & Walker, 2017 for a recent review and meta-analysis of the literature examining reactivity associated with making metacognitive judgements). Sub-group analysis has shown that when paired associates task word pairs consist of unrelated pairs or lists comprised of a mixture of related and unrelated pairs no reactivity is found in terms of impact on memory performance (Double, Birney, & Walker, 2017). In contrast, when tasks consist exclusively of related pairs there is moderate positive reactivity observed. Evidence indicates that older adults' learning from text is reduced by making metacognitive judgments (Stine-Morrow, Shake, Miles, & Noh, 2006).

Past studies examining reactivity associated with having participants make metacognitive judgements have focused on the effect of making judgements on task performance. However, it is possible that making judgements may alter frequency and content of metacognitive thinking without necessarily altering task performance. Additional studies are needed that examine the effect of requiring participants to make explicit metacognitive judgements on both task performance and participants' underlying thought patterns.

This type of reactivity could be problematic when the purpose of metacognition studies is to examine the correspondence between metacognitive judgments and behavior, as in studies that examine the relationship between metacognitive monitoring and self-regulation of study (e.g. Dunlosky & Connor, 1997; Dunlosky & Hertzog, 1997; Metcalfe & Finn, 2008). If requiring participants to make explicit judgments about task performance results in increased monitoring and, as a result, altered task performance, then it is difficult to know how generalizable the findings of metacognition studies are beyond the specific experimental setting.

The mind-wandering literature may provide a method for studying metacognitive monitoring, particularly the frequency of monitoring, in a less obtrusive way. As mentioned, research on mind-wandering often focuses on the content and nature of individuals' spontaneous thoughts. As such, methodological tools used to study mind-wandering may provide a tool for measuring spontaneous metacognitive thoughts without having participants make explicit performance judgments. While past studies of mind-wandering have relied on retrospective thought content questionnaires to assess individuals' self-generated intrusive thoughts while they complete cognitive tasks, many current studies instead rely on *online thought probes* to capture and measure mind-wandering experiences. Online thought probes may also be useful in identifying instances of spontaneous metacognitive monitoring.

Mind-Wandering Thought Probe Methodology

Mind-wandering if often assessed using online thought content probes that are embedded within a cognitive task of interest, often working memory span tasks or sustained attention tasks. Probes can vary in the types of 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 two to three minutes apart in order to allow the mind to begin wandering across task trials. An example of a typical thought probe appears below. Participants are asked "*What were you just thinking about*?" and respond by selecting one of the following options listed 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 provided in the online probe correspond to different types of thoughts the participants may have been experiencing during the task. Option (1) corresponds to being completely on-task and thinking only about responding appropriately to the task stimuli, option (2) corresponds to thoughts regarding task performance (*task-related interference*; TRI), 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.

Using online thought probes to capture different types of thoughts during a cognitive task is an approach used within the mind-wandering literature to measure the frequency of internally generated off-task thoughts individuals experience (i.e., mind-wandering) but it is also possible that this thought probe methodology can be used to measure the frequency of peoples' spontaneous metacognitive thoughts. Again, what is known as task-related interference within the mind-wandering literature involves off-task thoughts regarding one's task performance. This is likely to be analogous to metacognitive monitoring. Thought probes such as the one shown above provide a way for participants to indicate the amount of task-related interference or metacognitive monitoring they experience during a task.

Asking participants to make explicit metacognitive performance judgments may change the extent to which individuals engage in thoughts about their task performance. If making the typical, explicit metacognitive judgments such as JOLs and FOK judgments alters the frequency of thoughts about task performance, then different methods are needed to determine the frequency in which individuals spontaneously think about task performance during cognitive tasks. The more general thought probes used to study mind-wandering may be a suitable approach.

Although there has been concern that participants' (particularly older adults') thought reports may not be valid due to participants either being unaware of their mind-wandering or unwilling to report it, this does not seem to be the case. In one study examining mind-wandering and reading comprehension in younger and older adults, TUTs were associated with poorer reading comprehension, with no age differences

(Frank, Nara, Zavagnin, Touron, & Kane, 2015). Eye-movements were also measured, and individuals in both age groups made more regressions and engaged in more blinking when reporting TUTs than when they reported being on-task (Frank et al., 2015). These eye-tracking results along with the reading comprehension results do not support the hypothesis that older adults are unaware of their mind-wandering experiences and misclassify TUTs as being on-task. Instead, both older and younger adults seem to be aware of and willing to report their mind-wandering experiences. In addition, when younger and older adults provide pre-task predictions regarding how much expect to mind-wander during the subsequent task, older adults predict that they will spend more time off-task than younger adults predict (Frank et al., 2015), even though older adults later mind-wander less than younger adults. This suggests that older adults are *not* unwilling to report mind-wandering and it is therefore unlikely that older adults under-report their mind-wandering experiences due to fear of being stigmatized.

Research Aims

The dissertation study attempts to answer some of these questions. In this study, thought content probes from the mind-wandering literature were used investigate the frequency of metacognitive monitoring in younger and older adults. A primary focus of the study is to investigate whether the traditional methods used to measure metacognition (e.g. having participants make explicit metacognitive judgments) are reactive in terms of whether they alter monitoring frequency in participants who are required to provide these judgments. Additionally, the relationship between monitoring accuracy and monitoring frequency in both younger and older adults was investigated within the study. This study used a paired associates learning task to better connect to the existing metacognition literature.

Before describing the primary dissertation study, I will discuss a series of initial studies. As previously mentioned within the discussion of TRI, the specific content of TRI experiences remains unknown and it is unclear whether TRI can be thought of as being analogous to metacognitive monitoring. Studies 1 and 3 were designed to gain a better understanding of what types of metacognitive thoughts TRI encompasses. Studies 2 and 4 were also conducted to study the specific content of individuals' TRI experiences. An additional goal of studies 2 and 4 were to determine whether having participants complete more detailed thought content probes than those traditionally used in mind-wandering studies would lead to reactivity issues and under-reporting or over-reporting of TRI. Addressing this question allows us to determine which type of thought probe is best suited to measuring the amount of metacognitive monitoring individuals experience during an ongoing cognitive task. Below, I will discuss the main findings of studies 1 through 4, and describe in more detail the thought content probe proposed for Study 5.

Because TRI is typically reported infrequently by younger adults, we used a stereotype threat manipulation in studies 1 and 2 to elicit TRI our younger adult sample. Because TRI was examined in younger adults under stereotype threat in Studies 1 and 2 and because the older adult participants used in studies 3 and 4 were not exposed to a stereotype threat manipulation, the younger adult studies (1 and 2) and the older adult studies (3 and 4) will be discussed separately from each other. Additional figures that show overall TRI, proportions of the different subtypes of approach-based TRI, and proportions of the different subtypes of evaluative TRI with younger and older adult data plotted on the same figures are presented in Appendices O, P, and Q, respectively.

CHAPTER II

YOUNGER ADULT STUDY 1 METHODS

Study 1 was designed to gain a better understanding of what participants think about when they report TRI. The thought probes used in our past studies included two categories of TRI: thinking about task strategy or task approach (which we will call *approach-based TRI*) and thinking about task performance (which we call *evaluative TRI*). However, by using more open-ended thought probes, we can garner information regarding the specific content of individuals' mind-wandering experiences. In Study 1, we had a small sample of younger adults complete a task containing open-ended thought content probes.

Participants

Ten female younger adults aged 18-21 (M_{AGE} = 19.10) participated in the study. Participants were recruited via the UNCG Psychology SONA system and received course credit for their involvement in the study.

Materials and Procedure

Again, because TRI is typically reported infrequently by younger adults, we used a stereotype threat manipulation to elicit TRI our younger adult sample. To induce stereotype threat, we used a procedure previously used by our lab to elevate TRI in female undergraduates (Jordano & Touron, 2017b). After arriving to the lab and providing consent, participants completed a few basic demographics questions.

Participants were told that they were about to complete a task designed to measure “quantitative capacity” in order to collect normative data about college students, and that this task had revealed gender differences in the past. Additionally, participants were tested by a male experimenter and were always tested in a room along with either a male participant or a male confederate.

Participants completed an automated version of the OSPAN, during which participants alternated between verifying mathematical equations and remembering letters in serial order. The OSPAN consisted of 81 trials with set size varying between three, four, and five math problems to be verified and letters to be recalled. Participants practiced the math verification portion of the task, the letter recall portion of the task, and both the math verification and letter recall portions together along with an example of the thought probe before data collection actually began. During the task, participants received feedback regarding their trial-level letter recall accuracy, their trial-level math verification accuracy, and their cumulative math verification accuracy.

Nine thought probes were embedded within the OSPAN task, occurring at unpredictable intervals. Probes appeared at various points within an OSPAN trial. Probes appeared: (1) after participants saw the final mathematical equation within a trial, but before they verified the answer for that equation, (2) after the final mathematical equation within a trial had been verified, but before the final letter to be recalled was presented, and (3) at the end of a trial after participants saw the final letter to be recalled. Three probes appeared during trials containing set sizes of three, three probes appeared during trials containing set sizes of four, and three probes appeared during trials containing set sizes of five. These thought probes were entirely open ended and asked participants to type in a response to the question, "*What were you just thinking about*?" A space was provided below this prompt for participants to type in their response.

Participants next completed various post-task questionnaires. These included the Thought Content components of the Dundee Stress State Questionnaire (DSSQ; Matthews et al., 1999; see Appendix A) which was used as an additional retrospective measure of mind-wandering. The DSSQ contains eight questions assessing TRI (e.g., "I thought about how I should work more carefully") and eight questions assessing TUT (e.g., "I thought about something that happened to me earlier"). The Motivation component of the DSSQ, which contains 15 questions related to motivation level, was used as a measure of participant motivation (see Appendix B). Participants also responded to several additional Likert-scale questions including questions about perceived task difficulty, task focus, fatigue, and stress (see Appendix C for a full list of single-item post-task questions; see Appendix D for younger adult Study 1 post-task descriptive statistics). Results for Study 1 are presented alongside of the results for Study 2 to make it easier to see how *probe type* influences the frequency of TRI and the specific content of participants' TRI experiences.

CHAPTER III

YOUNGER ADULT STUDY 2 METHODS

In addition to examining the specific content of individuals' TRI experiences, a goal of the younger adult Study 2 was to determine whether having participants respond to more detailed thought probes that include follow-up questions regarding one's thought content would lead to reactivity. If participants must complete additional follow-up questions regarding thought content if they indicate experiencing TRI and TUTs, participants might be more likely to indicate that they were on-task. To examine this issue, younger adults were randomly assigned to one of two conditions. Participants in the control condition completed a cognitive task that contained typical thought content probes. Participants in the experimental condition completed the same experimental task, but responded to thought content probes that branched and asked additional, follow-up questions regarding thought content.

Participants

Sixty female younger adults aged 18-22 (M_{AGE} = 19.58) participated in the study. Participants were recruited via the UNCG Psychology SONA system and received course credit for their involvement in the study. Participants were randomly assigned to either a standard thought probe condition or a branching thought probe condition.

Materials and Procedure

Again, we used a stereotype threat manipulation (Jordano & Touron, 2017b) to elicit stereotype threat and TRI in our younger adults. The stereotype threat manipulation used in Study 2 was identical to the one used in Study 1. Participants completed an automated version of the OSPAN. The OSPAN used was the same as the one used in Study 1, with the exception of the types of thought probes embedded within. Thought probe placement was identical to Study 1. Participants in the standard probe condition responded to the standard thought probe that has been used in our past mind-wandering studies. Standard probe participants were asked, "*What were you just thinking about*?" and saw a corresponding thought content probe. Figure 1 below shows the different levels of thought probes participants could respond to. All participants saw and responded to Level 1 (the standard thought probe) and, depending on how they responded to Level 1, participants in the branching probe group saw and responded to an additional Level 2 thought probe. Participants next completed various post-task questionnaires. Participants completed all the same post-task measures that were included in Study 1 (see Appendix C for a full list of single-item post-task questions; see Appendix E for younger adult Study 1 post-task descriptive statistics)

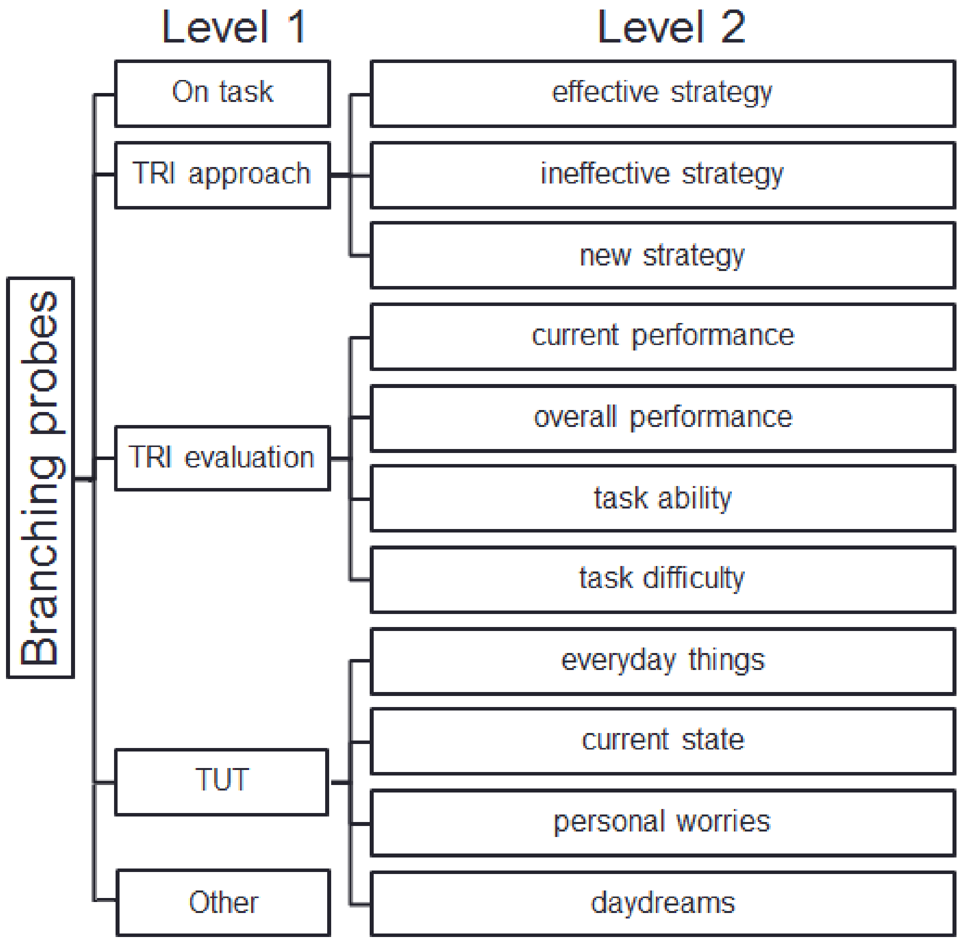


Figure 1. Thought Probe Levels Used in Study 2 and Study 4.

CHAPTER IV

YOUNGER ADULT STUDIES RESULT AND DISCUSSION

Responses to the open-ended thought probes in Study 1 were coded into different mind-wandering categories. Responses could be coded as being: on-task, approach-based TRI, evaluative TRI, TUTs about everyday things, TUTs about current state, TUTs about personal worries, TUTs about daydreams, or other. Open-ended thought probe responses were coded into these categories by three independent raters, with near perfect agreement. (ICC = .96). There were no open-ended probe responses that did not fit into the thought categories listed above.

Participants in Study 1 reported mean proportions of TRI, approach-based TRI, and evaluative TRI that were comparable to what was reported by participants in Study 2 (see Figure 2). Proportion of approach-based TRI is measured as the amount of reported approach-based TRI out of total amount of thought reports. Proportion of evaluative TRI is measured as the amount of reported evaluative TRI out of total amount of thought reports. For Study 2, we found no effect of thought probe type on mean proportion of overall TRI, F (1, 58) = .114, p = .736, indicating that having participants answer additional questions about specific TRI content did not lead to reactivity or over or under-reporting of TRI. There was no condition x TRI subtype interaction. Participants reported similar proportions of approach-based and evaluative-TRI regardless of the type of probe they were required to response to. Participants in Study 2 did report more TRI approach-based TRI compared to evaluative TRI, F (1, 58) = 6.65, p = .012.

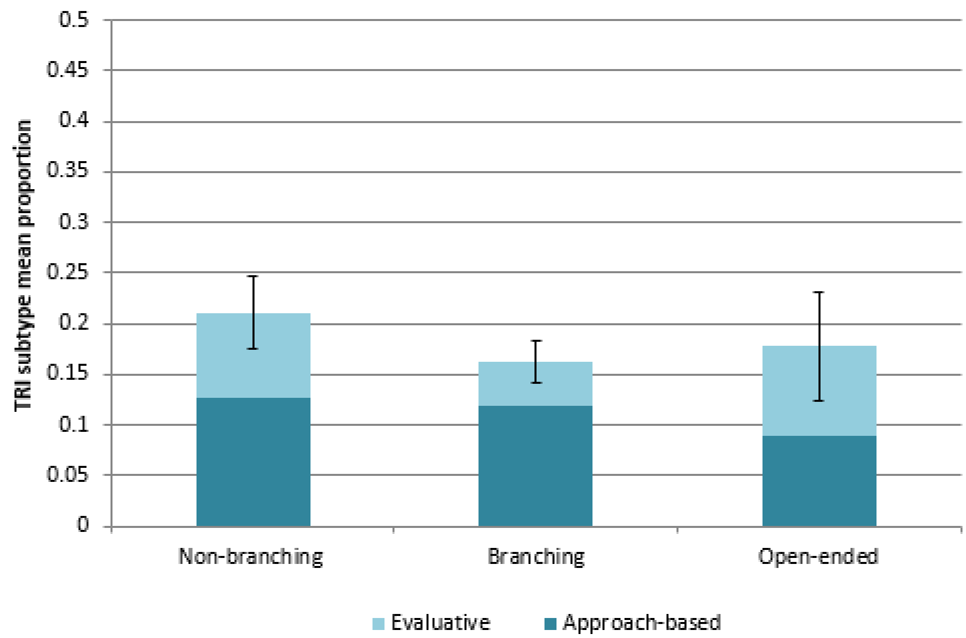


Figure 2. TRI for Studies 1-2.

In these studies, we broke approach-based and evaluative TRI into different content subcategories. If participants reported engaging in approach-based TRI, they reported whether they were thinking about: (1) whether their current strategy was effective (2) whether their current strategy was ineffective or (3) a new strategy that could be implemented on future OSPAN trials (see Figure 3). Values in Figure 3 reflect the amount of each approach-based TRI subcategory *as a proportion of the total amount of reported approach-based TRI.*

Participants in Study 1 as well as both conditions of Study 2 were more likely to have thoughts where they evaluated their current task strategy than they were to think about new strategies that could be used on future trials (see Figure 3). In our original conceptualization of these thought types, we labeled approach-based TRI as “proactive TRI” because we believed participants might think about strategies they could use on future trials to improve task performance. While participants did report thinking about task approach or strategy, they were not thinking about specific ways to improve performance on the task. Rather, they were evaluating the effectiveness of the strategy they were currently using. Evaluating the effectiveness of the task strategy one is currently using could be considered a form of metacognitive monitoring rather than an example of proactive metacognitive control.

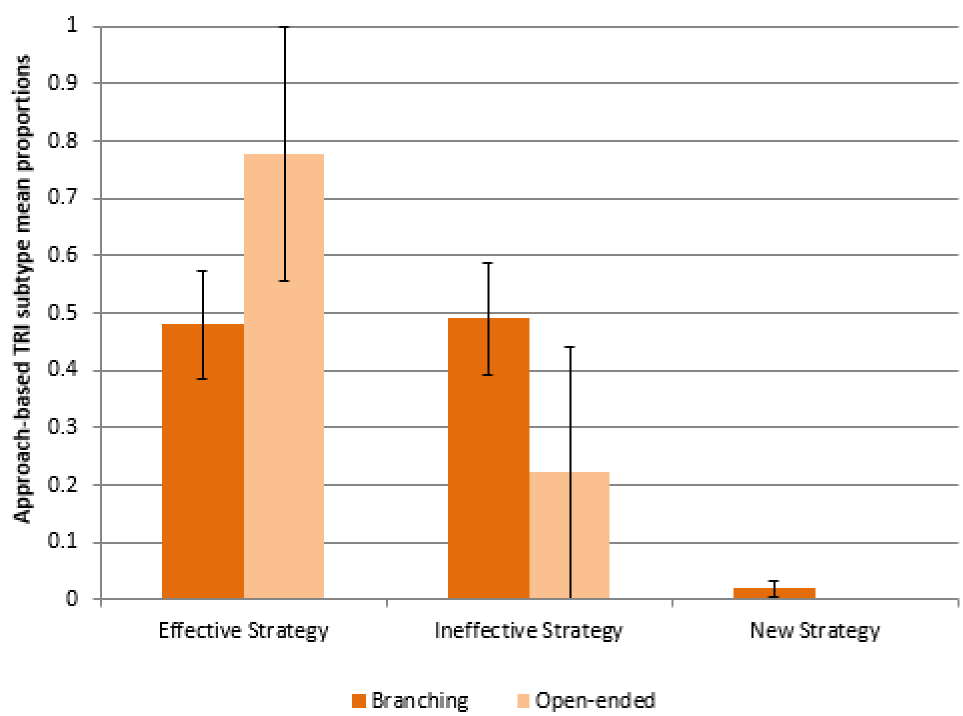


Figure 3. Approach-Based TRI for Studies 1-2.

If participants reported engaging in evaluative TRI, they reported whether they were thinking about: (1) their current, trial-level task performance (2) their overall task performance on the OSPAN (3) their general task ability, or (4) the difficulty of the task (see Figure 4). Values in Figure 4 reflect the amount of each evaluative TRI subcategory *as a proportion of the total amount of reported evaluative TRI*. Participants in Study 1 and within both condition of Study 2 reported thinking about current performance, overall performance, and their general ability to do the task well. No participants reported thinking specifically about the difficulty of the task (see Figure 3). This pattern may provide valuable insight into the impact of stereotype manipulations on thoughts during a task, with dispositional rather than situational attributions of performance concerns (Schmader, 2010).

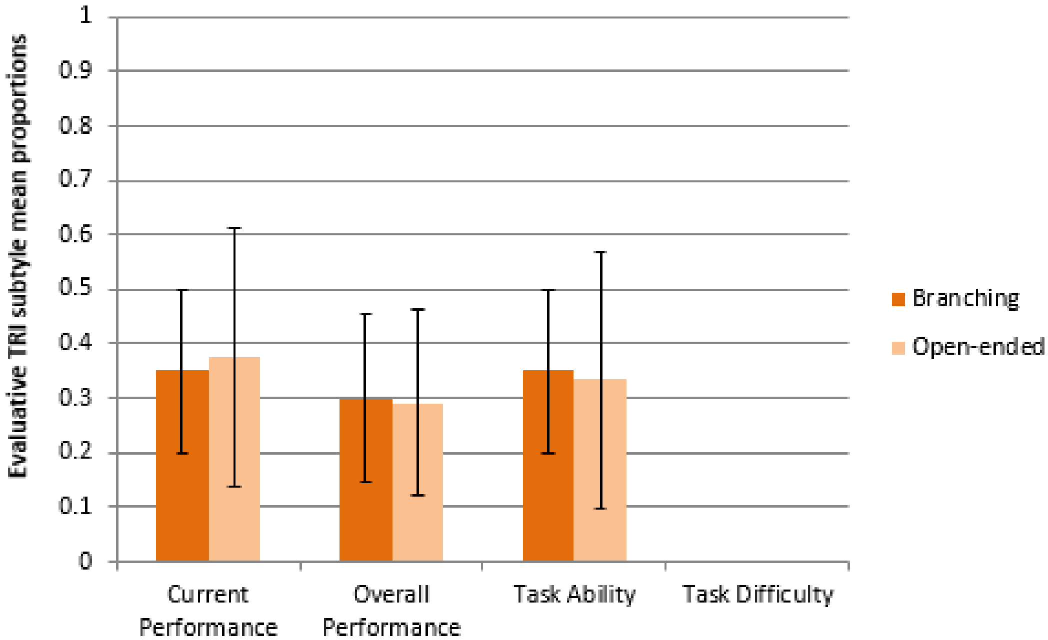


Figure 4. Evaluative TRI for Studies 1-2.

The results of these initial studies suggest that mind-wandering probes do not alter underlying cognitive processes. Having participants answer follow-up questions about the specific content of their TRI (metacognitive) experiences in Study 2 does not lead to underreporting or over-reporting of TRI. Additionally, is seems that participants engage in metacognitive monitoring processes when they report approach-based and evaluative TRI. When participants reported evaluative TRI, they often thought about their task performance and ability to perform the task well. When participants reported approach-based TRI, they monitored the effectiveness of their current task strategy. Therefore, TRI seems to be analogous to metacognitive monitoring, at least in the present OSPAN task.

Given these findings, mind-wandering thought probes including TRI as a response option may be useful in studying the frequency of metacognitive monitoring.

Since having participants answer follow-up questions about their TRI experiences does not seem to lead to under or over-reporting of TRI, future studies examining monitoring frequency could use the more nuanced thought probes developed in Study 2.

CHAPTER V

OLDER ADULT STUDY 1 METHODS

Using more detailed thought probes than the ones used in previous studies of mind-wandering, Studies 1 and 2 provided us a better understanding of what individuals think about when they report TRI. When younger adults in Study 1 and Study 2 reported TRI, they primarily engaged in monitoring of their task performance as well as monitoring of the effectiveness of their task strategies. These younger adult studies suggest that TRI is analogous to metacognitive monitoring and, as such, we might be able to use online, branching thought probes such as the one used in Study 2 to capture and study the frequency of metacognitive monitoring.   Before using this type of thought probe to study monitoring frequency in older adults, we will attempt to replicate the results of the younger adult Study 1 and Study 2 with an older adult sample.to verify that TRI experienced by older adults is analogous to metacognitive monitoring.

Participants

Ten older adults (aged 60-75, $M_{\text{AGE}} = 66.58$) participated in the study. Participants were recruited from the Adult Cognition Lab database of community older adult volunteers. Older adults were paid $10 per hour for participation and participation lasted two hours. All participants were screened for near visual acuity of 20/50 or better and participants with a history of stroke or dementia were excluded from participating in the study.

Materials and Procedure

After arriving to the lab and providing informed consent, participants completed a few basic demographics questions, in addition to measures of near visual acuity, perceptual speed and crystallized intelligence. Participants completed the same automated version of the OSPAN used in the younger adult studies. As in the younger adult Study 1, participants responded to nine, open-ended thought probes embedded within the OSPAN task. Responses to these open-ended thought probes were coded into the same mind-wandering categories that were used in the younger adult Study 1. Participants also completed the same post-task questionnaires that were included in the younger adult studies.

CHAPTER VI

OLDER ADULT STUDY 2 METHODS

Participants

Sixty older adults (aged 60-75, M_{AGE} = 65.97) participated in the study. Participants were recruited from the Adult Cognition Lab database of community older adult volunteers. Older adults were paid $10 per hour for participation and participation lasted two hours. All participants were screened for near visual acuity of 20/50 or better and participants with a history of stroke or dementia were excluded from participating in the study.

Materials and Procedure

After arriving to the lab and providing informed consent, participants completed a few basic demographics questions, in addition to measures of near visual acuity, perceptual speed and crystallized intelligence. Participants were randomly assigned to either the standard thought probe or branching thought probe condition. Participants next completed the same automated version of the OSPAN used in the younger adult Study 2. Participants completed the same post-task questionnaires used in the younger adult initial studies.

CHAPTER VII

OLDER ADULT STUDIES RESULTS AND DISCUSSION

Responses to the open-ended thought probes in Study 3 were coded into different mind-wandering categories. Responses could be coded as being: on-task, approach-based TRI, evaluative TRI, TUTs about everyday things, TUTs about current state, TUTs about personal worries, TUTs about daydreams, or other. Open-ended thought probe responses were coded into these categories by three independent raters, with near perfect agreement (ICC = .97). There were no open-ended probe responses that did not fit into the thought categories listed above.

Participants in older adult Study 3 reported proportions of TRI, approach-based TRI, and evaluative TRI that were comparable to what was reported by participants in Study 4 (see Figure 5). Across older adult Study 3 and Study 4, we found no effect of thought probe type on overall TRI, F (1, 69) = 2.173, p = .122, indicating that having participants answer additional questions about specific TRI content did not lead to reactivity. Proportion of approach-based TRI was operationalized as the amount of reported approach-based TRI out of total amount of thought reports. Proportion of evaluative TRI was operationalized as the amount of reported evaluative TRI out of total amount of thought reports. Probe type did not affect the proportion of approach-based TRI, F (1, 69) = 1.802, p = .173. While, participants in Study 3 did report more evaluative TRI than participants in Study 4, this difference in proportion of evaluative TRI was not significant, F (1, 69) = 2.860, p = .064.

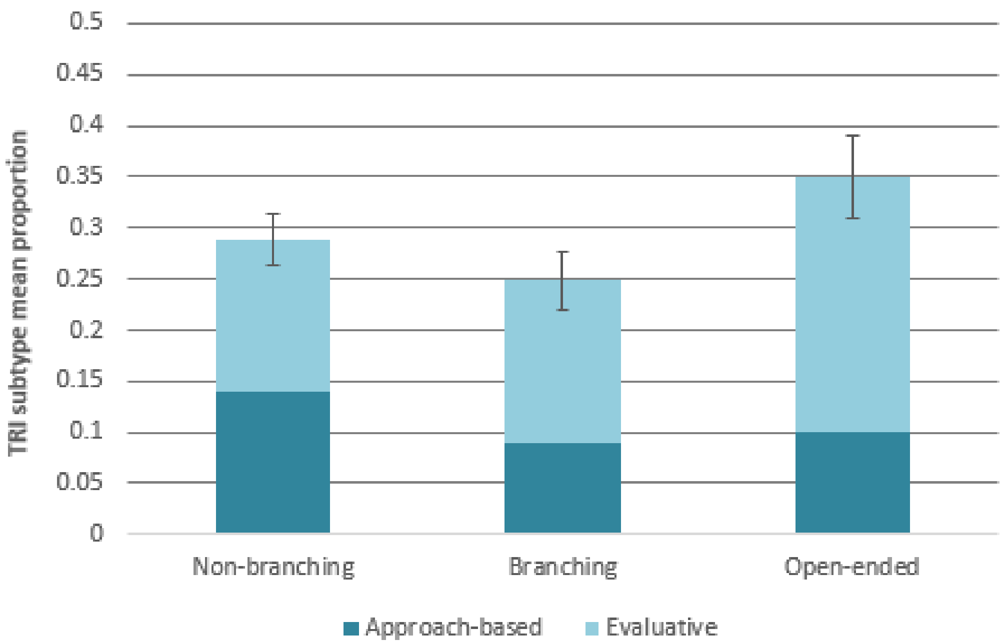


Figure 5. TRI for Studies 3-4.

Values in Figure 6 reflect the amount of each approach-based TRI subcategory *as a proportion of the total amount of reported approach-based TRI.* Again, in previous studies within our lab (Jordano & Touron, 2017a; Jordano & Touron, 2017b), we called approach-based TRI as “proactive TRI” because we believed participants might think about strategies they could use on future trials. While participants did report thinking about task approach or strategy, they generally evaluated the effectiveness of their current task strategy rather than thinking about new and potentially more effective strategies. Evaluating the effectiveness or ineffectiveness of task strategy could be considered a form of metacognitive monitoring rather than an example of proactive, metacognitive control. Additionally, older adults in Study 3 reported significantly more TRI about the effectiveness of currently used strategies than older adults in Study 4, t (38) = 2.222, p = .039, d = .670.

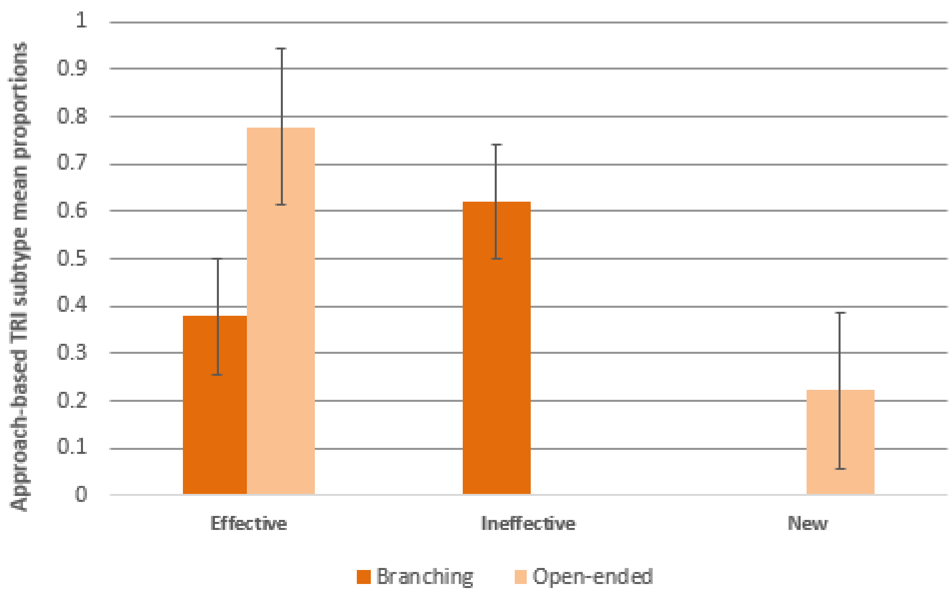


Figure 6. Approach-Based TRI for Studies 3-4.

If participants reported evaluative TRI, they reported whether they were thinking about: (1) their current, trial-level task performance (2) their overall task performance on the OSPAN (3) their general task ability, or (4) the difficulty of the task (see Figure 7). Values in Figure 7 reflect the amount of each evaluative TRI subcategory *as a proportion of the total amount of reported evaluative TRI.* Unlike the younger adults in Study 1, older adults did report thinking specifically about the difficulty of the task. Across the two older adult studies there were no significant differences in the amount of reported TRI about current performance, task ability, and task difficulty (all p-values > .05 for t-test focused comparisons). Older adults in Study 4, who responded to a thought probe containing different thought categories that they must choose from, reported significantly more TRI regarding their overall task performance than older adults in Study 3, t (38) = 2.830, p = .009, d = 1.055. This finding suggests that older adults are not as likely to reflect upon their global task performance when they are *not* presented with this thought category on a thought probe. It is worth noting that this was not found in younger adults (see Figure 3 above). Younger adults were equally as likely to report reflecting on their overall performance regardless of whether they saw this thought type option on the probe. Additional work is needed to determine why viewing thought probes that explicitly list thinking about overall task performance as a thought option alters monitoring in older but not younger adults.

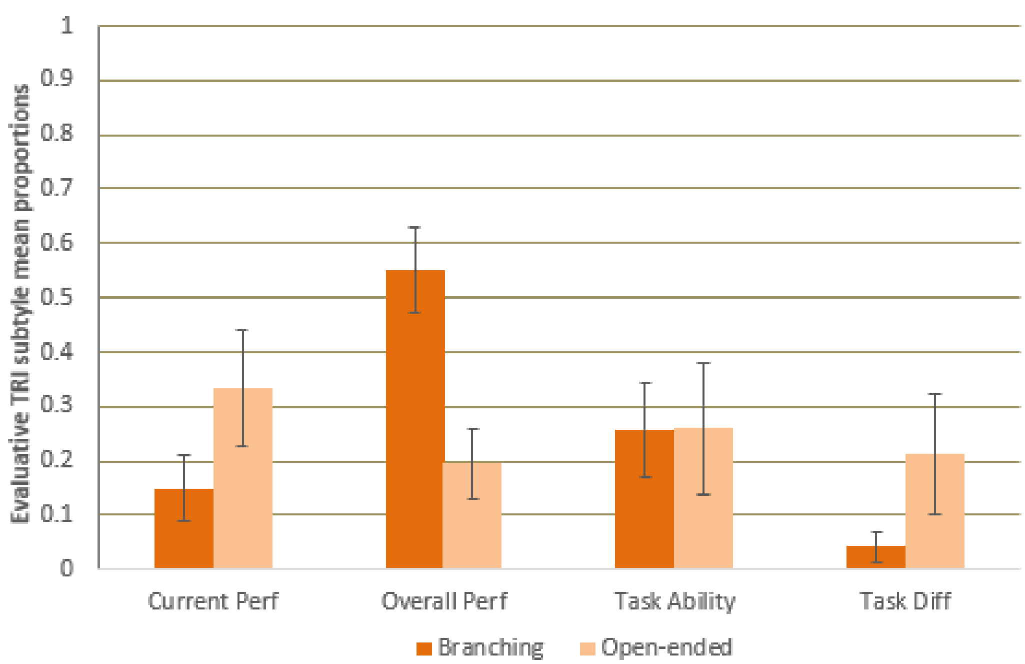


Figure 7. Evaluative TRI for Studies 3-4.

As with the younger adults Study 1 and 2, the results of the older adult Study 3 and Study 4 suggest that participants engage in monitoring processes when they report both evaluative TRI and approach-based TRI. Like younger adults, when older adults report approach-based TRI, they report monitoring and evaluating the effectiveness or ineffectiveness of their current strategy. Therefore, TRI seems to be analogous to metacognitive monitoring in both younger and older. Given this finding, mind-wandering thought probes that include TRI may be useful in studying the frequency of metacognitive monitoring in older adults.

Initially, we believed that if we could replicate the results of younger adult Study 1 and 2 in an older adult sample, we would use the branching thought probes used in Study 2 and Study 4 to investigate monitoring frequency in Study 5. Although type of thought probe (open-ended, standard, or branching) did not affect the overall amount of TRI reported by older adults, thought probe type did affect reported amount of TRI about strategy effectiveness, strategy ineffectiveness, new strategies, current performance, and task difficulty. To avoid potential under- or over-reporting of TRI in older adults, and to get a better idea of the content of participants' TRI during the paired associates task used in Study 5, we will use open-ended thought probes rather than branching thought probes.

CHAPTER VIII

STUDY 5 METHODS

As previously indicated, two primary goals of Study 5 are to identify age-related patterns in the frequency of metacognitive monitoring, to determine whether making monitoring judgments impacts the frequency of monitoring, and to investigate the relationship between monitoring *frequency* and monitoring *accuracy*. Spontaneous thoughts are likely to vary markedly with different kinds of tasks; we were particularly interested in looking at these phenomena in the context of a paired associates learning task given the ubiquity of these tasks in the metamemory literature. To address these questions, younger and older adults completed a task in which they were required to study stimuli for a later memory test (the paired associates task). During this task, participants answered thought content probes identical to the ones used in Study 1 and Study 3 described above and, in one condition, were required to make metacognitive judgments regarding their learning of the task stimuli.

A sample of younger adults and a sample of older adults were randomly assigned to one of two conditions. Participants in the probe only (control) condition completed a learning task that contained only thought content probes. Participants in the JOL + probe (experimental) condition completed the same learning task, but with both thought content probes and questions asking participants to make metacognitive judgments embedded within. This design allowed us to determine how having participants make explicit metacognitive judgments alters participants' propensity to monitor the task. Additionally, we conducted an age group comparison to determine age-related patterns in monitoring frequency.

Younger and older adults within the JOL + probe condition were also asked to provide metacognitive judgments of learning (JOLs) during the learning task. Having participants within the JOL + probe group make JOLs in addition to responding to thought probes allowed us to measure the association between monitoring frequency (as measured by the thought content probes) and monitoring accuracy, which were obtained using these participants' JOLs.

Participants

An a priori power analysis in G*Power (Faul, Erdfelder, Buchner, & Lang, 2007) suggested that a sample of 210 participants (53 participants per age x condition cell) should be sufficient to achieve 95% power to detect a medium effect size for a 2 (age: young, old) x 2 (condition: probe only, JOL + probe) factorial ANOVA with age and condition as between-subject variables. This sample size calculation was conducted using the traditional .05 criterion of statistical significance. Given this, we will test a total of 106 younger adults (aged 18-25) and 106 older adults (aged 60-75).

One hundred and six younger (aged 18-25, M_{AGE} =19.60 years, SD_{AGE} = 1.14 years) and 106 older adults (aged 60-75, M_{AGE} = 67.51 years, SD_{AGE} = 3.89years) were tested. Younger adults were recruited from the UNCG Psychology SONA system and older adults were recruited from the Adult Cognition Lab database of community older adult volunteers. Older adults were paid $10 per hour for participation. All participants were screened for near visual acuity of 20/50 or better and participants with a history of stroke or dementia were excluded from participating in this study. Older adults (M = 2.148, SE = .237) reported taking more medications than younger adults (M = .867, SE = .184), t (210) = 4.267, p < .001, d = .581. Older adults (M = 23.367, SE = 1.177) also scored lower on processing speed than younger adults (M = 35.833, SE = 13.021), t (210) = 6.998, p < .001, d = .926. Finally, older adults (M = 23.033, SE = 1.253) had higher vocabulary scores than YAs (M = 13.733, SE = .663), t (210) = 6.560, p < .001, d = .893. These patterns are all consistent with those typically obtained in cognitive aging research. Importantly, there were no condition differences between participant groups in terms of age, number of medications, processing speed, and vocabulary (all p-values > .05).

Materials

Paired associates learning task. As mentioned within the introduction, the paired associates task is commonly used within the study of metacognitive accuracy and control (Arbuckle & Cuddy, 1969; Leonesio & Nelson, 1990; Nelson & Dunlosky, 1991; Dunlosky & Nelson, 1994). In these past metacognition studies that have used the paired associates task, participants are required to provide metacognitive judgments (JOLs) for the study items. Because the paired associates task is so commonly used to study metamemory, we use it in the current study to investigate whether requiring participants make JOLs influences their propensity to engage in monitoring during this task.

Participants in the current study completed the paired associates learning task programmed in E-Prime 2. Participants completed two blocks of the paired associates task, with Block 1 and Block 2 each consisting of a study phase followed by a cued recall test. The study phases in Block 1 and Block 2 each included a total of 60 word pairs that were studied by participants for a later memory test. Prior to completing Block 1 participants studied and recalled an additional six items that served as practice items.

The word pairs used included words that varied in their level of intrinsic association. Two lists of 60 word pairs were constructed so as to include 20 strongly associated, 20 weakly associated, and 20 unrelated pairs. Associative strength is defined here as the probability of producing the second word of the word pair as an associate of the first word in the word pair (see Koriat, 1981). Strongly associated pairs had probabilities of association between .400 and .750, weakly associated pairs had probabilities of association no greater than .040, and unrelated pairs had probabilities of association of zero (see Castel, McCabe, & Roediger, 2007). All items used in the paired associates task were nouns taken from the Nelson, McEvoy, and Schreiber free association norms (2004). Presentation order was randomized during both practice and the experimental task, with a one second inter-stimulus interval between the presentations of each word pair.

Each word pair appeared in the center of the computer screen for a total of 10 seconds. In both the probe only and JOL + probe conditions, thought probes appeared approximately every 90 seconds. On probe trials, participants studied the word pair for eight seconds before the task was interrupted and the thought probe appeared on the computer screen. Due to the possibility of forced-choice thought probes leading to over or under-reporting of TRI, open-ended thought probes identical to the ones used in Studies 1 and 3 were used. Participants used their computer keyboard to type in whatever they had been thinking about immediately before the thought probe appeared. Participants responded to a total of 20 open-ended thought probes. Ten probes appeared during the study phase of Block 1 and 10 probes appeared during the study phase of Block 2.

In the JOL + probe condition, after studying each word pair for eight seconds, participants saw a JOL screen. On the JOL screens, the stimulus word of the pair that was just studied appeared on the screen (e.g., SUGAR-???) along with the instructions: "How confident are you that in about 20 minutes from now you will be able to recall the second word of the item when prompted with the first word?" JOL + probe condition participants were instructed to type in a percentage from 0% to 100%, indicating the likelihood that they will be able to successfully recall the second word in the word pair they just studied when cued with the first word (Schraw, 2009). Participants made their JOLs by typing in their chosen percentage. After the JOL has been entered, the next word pair was displayed for study. In the JOL + probe condition, JOL screens did not appear on the trials that contained thought probes. Participants in the probe only group were not required to make JOL responses on any trials, and the study pair was instead presented for the full 10 second duration.

Paired associates cued recall task. As mentioned above, the paired associates task was divided into two blocks, each consisting of a study phase followed by a cued recall test. During this test, participants were given the first word of each pair they had studied in the preceding study phase and were prompted to type in the second word of the pair. The order of item presentation was again randomized.

Scoring. The thought content probes answered by participants within both conditions were used to measure monitoring frequency. Because we believe metacognitive monitoring to be analogous to task-related interference (TRI), mean proportion of overall TRI served as our measure of monitoring frequency. Relative accuracy of JOLs was operationalized as the intraindividual gamma correlation (Nelson, 1984) between JOLs and recall test performance. Again, these can be interpreted as the conditional probability that recall of Item A is greater than that of Item B, given that Item A was provided with a higher JOL than Item B was.

Paired associates strategy reports. Following the recall test in Block 2, the item pairs from Block 2 were shown in their original study order, and individuals reported which strategy (if any) they used to learn each item. Participants were given brief descriptions of different mnemonic strategies so they could provide item-level strategy reports (Dunlosky & Hertzog, 1998; Hertzog, Sinclair, & Dunlosky, 2010). Participants did so by selecting one of six options: 1 – rote repetition, 2 – interactive imagery, 3 – sentence generation, 4 – some other strategy, 5 – no strategy, 6 – tried to use a strategy but ran out of time.

Post-task measures. In addition to the various pre-task questionnaires, the experimental paired associates task, and the follow-up cued recall test, participants completed a few computerized post-task questionnaires. Participants completed single-item questions about things such as perceived task difficulty, task interest, task motivation, and fatigue (see Appendix K). Participants also responded to a brief measure of everyday cognitive failures (The Cognitive Failures Questionnaire; Broadbent; Cooper, Fitzgerald, & Parkes, 1982; Appendix M). Finally, participants completed a brief measure of everyday monitoring experiences (adapted from the Dundee Stress State Questionnaire; Matthews et al., 2002; Appendix L), to obtain a rough measure of participants' propensity to engage in metacognitive monitoring outside of the lab.

Procedures

All measures were collected during one session. Sessions lasted approximately one and one-half hour for younger adults and approximately two hours for older adults. The study layout and specific components of the paired associates task (for the JOL + probe condition) is presented in Figure 6 below.

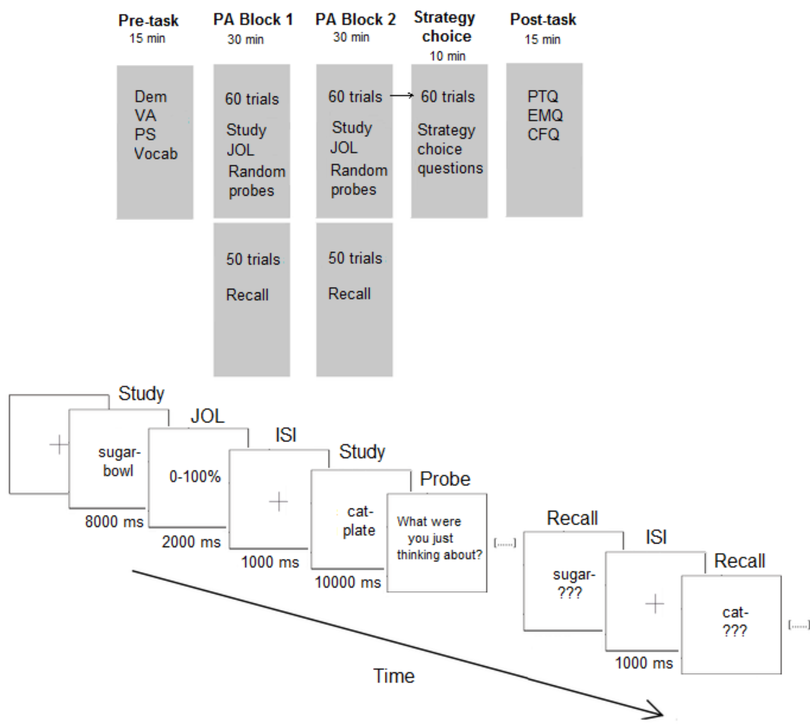


Figure 8. Layout of Study 5.

Note. The layout of the overall experiment (TOP) and the Paired Associates (PA) task for the *JOL + probe condition* (BOTTOM). TOP: Dem = demographics, VA = visual acuity, PS = processing speed, Vocab = vocabulary, PTQ = single-item questions, EMQ = everyday measure of monitoring, CFQ = cognitive failures questionnaire. BOTTOM: For each PA Block in the JOL + probe condition, each of 60 pairs were studied for 8 seconds, followed by a 2 second JOL, and then a 1 second ISI. After 10 of the pairs (distributed and pseudo-random), a thought probe occurred instead of the JOL screen. Following the study phase, participants performed a cued recall task for the 50 pairs that were not followed by thought probes; the test for each pair will be followed by a 1 second ISI. Probe only group participants completed the same task with the following exception; rather than making JOLs, study instead continued for the full 10 second.

CHAPTER IX

STUDY 5 RESULTS

The following analyses were conducted to address our questions of interest. First, we report the results of age (young, old) by condition (probe only, JOL + probe) ANOVAs on rates of TRI and recall accuracy. For the JOL + probe condition only, we examined age differences in JOLs and relative and absolute metacognitive monitoring accuracy. We also conducted age by condition ANOVAs for strategy use, TUT rates, and post-task survey responses. Lastly, we examined correlations between monitoring frequency and monitoring accuracy.

Participants completed two blocks of the paired associates task with each block consisting of 60 word pairs and 10 thought probes. Participants completed the recall test for Block 1 before they began Block 2 of the task. While participants did not receive explicit feedback regarding their recall performance it is still possible that completing the recall task before Block 2 affected the frequency of metacognitive monitoring, accuracy of metacognitive monitoring, and eventual paired associates recall performance for Block 2. Because of this, we conducted post-hoc analyses examining the effect of paired associates task *block* on our dependent variables. We report the results of age (young, old) by condition (probe only, JOL + probe) by block (Block 1, Block 2) mixed ANOVAs on rates of TRI and recall accuracy. For the JOL + probe condition only, we examined age and block differences in JOLs and relative and absolute metacognitive monitoring accuracy. We also conducted age by condition by block mixed ANOVAs for TUT rates. If applicable, main effects were followed-up by pairwise comparisons with Bonferroni correction.

Additional post-hoc analyses were conducted to examine the effect of paired associate *item relatedness* on our dependent variables. We report the results of age (young, old) by condition (probe only, JOL + probe) by relatedness (strongly related, moderately related, unrelated) mixed ANOVAs on rates of TRI and recall accuracy. For the JOL + probe condition only, we examined age and relatedness differences in JOLs and relative and absolute metacognitive monitoring accuracy. We also conducted age by condition by relatedness ANOVAs for strategy use and TUT rates. Main effects were followed-up by pairwise comparisons with Bonferroni correction.

Task-Related Interference

Open-ended thought probe responses were coded into the thought categories by three independent raters, with near perfect agreement (ICC = .95). Monitoring frequency was measured using the TRI category from the open-ended thought content probes and is operationalized as the mean proportion of reported TRI about task-approach and task performance. In keeping with previous findings from the mind-wandering literature (McVay et al., 2013; Zavagnin et al., 2014; Frank et al., 2015; Jordano & Touron, 2017a), we anticipated a main effect of age on monitoring frequency.We expected that older adults would engage in more monitoring and would thus report a higher proportion of TRI than younger adults. As expected, the proportion of TRI was higher for older adults (*M* =.213, *SE* = .021) than it was for younger adults (*M* =.097, *SE* = .020), *F* (1, 215) = 16.664, $p < .001$, $\eta p2$ = .073).

We also expected a main effect of condition on monitoring frequency, with both younger and older adults in the JOL + probe group reporting more monitoring than younger and older adults in the probe only group, due to the presence of the JOL screens. Overall, participants in the JOL + probe l condition (*M* =.219, *SE* = .020) reported more TRI than those in the probe only condition (*M* = .092, *SE* = .020), *F* (1, 215) = 19.700, $p < .001$, $\eta p2$ = .085).

It was initially hypothesized that these main effects would be qualified by an age x condition interaction. We expected that the inclusion of JOL screens in the JOL + probe condition would *decrease* age-related differences in the amount of reported TRI by increasing TRI in younger adults who might not engage in monitoring otherwise. An age x condition interaction (*F* (1, 215) = 10.081, *p* = .002, $\eta p2$ = .045) was found. However, younger adults in the JOL + probe l condition (*M* = .115, *SE* = .022) did *not* report significantly more TRI than younger adults in the probe only condition (*M* = .078, *SE* = .017), *t* (104) = 1.322, *p* = .189, *d* = .181. Conversely, older adults in the JOL + probe condition (*M* = .322, *SE* = .046) reported more TRI than older adults in the probe only condition (*M* = .105, *SE* = .020), *t* (104) = 4.334, $p < .001$, *d* = .589.

While responding to JOL screens did not seem to effect younger adults' propensity to engage in metacognitive thinking, the inclusion in JOL screens did increase the propensity to engage in metacognitive thinking in older adults, who may already more likely to experience concerns regarding cognitive performance (Hertzog & Hultsch, 2000). This pattern is interesting, because it indicates that older adults, who are already more prone to monitoring, are also more likely to increase monitoring when the task encourages these monitoring behaviors. In contrast, task demands have less of an impact on monitoring frequency in younger adults.

We conducted additional post-hoc analyses examining (1) the reliability of our TRI measure across the two halves of the paired associates task and (2) the effect of block on TRI. Descriptive statistics for proportion of overall TRI broken down by age group, condition, and task block are presented in Appendix R. To examine the reliability of our TRI measure across Block 1 and Block 2 of the paired associates task we calculated split-half reliability of our TRI measure. The Spearman-Brown split-half reliability coefficient for overall proportion of reported TRI across the two task blocks was .87, indicating *adequate* reliability of the TRI measure (Hulin, Netemeyer, & Cudeck, 2001). There was a significant main effect of block on TRI, F (1, 212) = 32.531, $p < .001$, $\eta p2 = .133$. Participants reported significantly more TRI on Block 1 (M = .093,SE = .008) of the PA task than on Block 2 (M = .062, SE = .007) (p-value < .05). The block x age, block x condition, and block x age x condition interactions were all non-significant (all p-values > .05).

It has been found that task *duration* affects TRI (Smallwood, Obonsawin, & Reid, 2002) with participants reporting fewer stimulus-dependent thoughts (on-task thinking and task-related interference) and more stimulus-independent thoughts (TUTs) with increased task duration. It is possible that both younger and older adults became less focused on the task overtime, reporting fewer on-task thoughts and TRI and more TUTs on Block 2 compared to Block 1.

Descriptive statistics for proportion of overall TRI broken down by age group, condition, and paired associates item relatedness are presented in Appendix S. A main effect of relatedness was found, F (1, 212) = 26.131, $p < .001$, $\eta p2 = .110$. Pairwise comparisons showed that participants reported significantly more TRI on probes immediately following an unrelated word pair ($M = .083$, $SE = .008$) compared to probes immediately following moderately related ($M = .059$, $SE = .007$) or strongly related $M = .041$, $SE = .005$) word pair (p-value < .05). TRI did not differ for probes that followed strongly related word pairs compared to probes that followed moderately related word pairs (p-value > .05). There were no significant age x relatedness, condition x relatedness, or age x condition x relatedness interactions (all p-values > .05).

Content of TRI

As in our past studies, we separated TRI into different sub-types, including monitoring of task performance (which we label "evaluative TRI") and mind-wandering about task strategy or task approach (which we label "approach-based TRI") (see Figure 8 below). In this study, we have also included a third category of TRI (which we label "other TRI") which includes task-related thoughts that do not fall neatly into the other two TRI categories. Examples of "other TRI" that participants reported include wondering how long the task will take and wondering when the memory test was going to start.

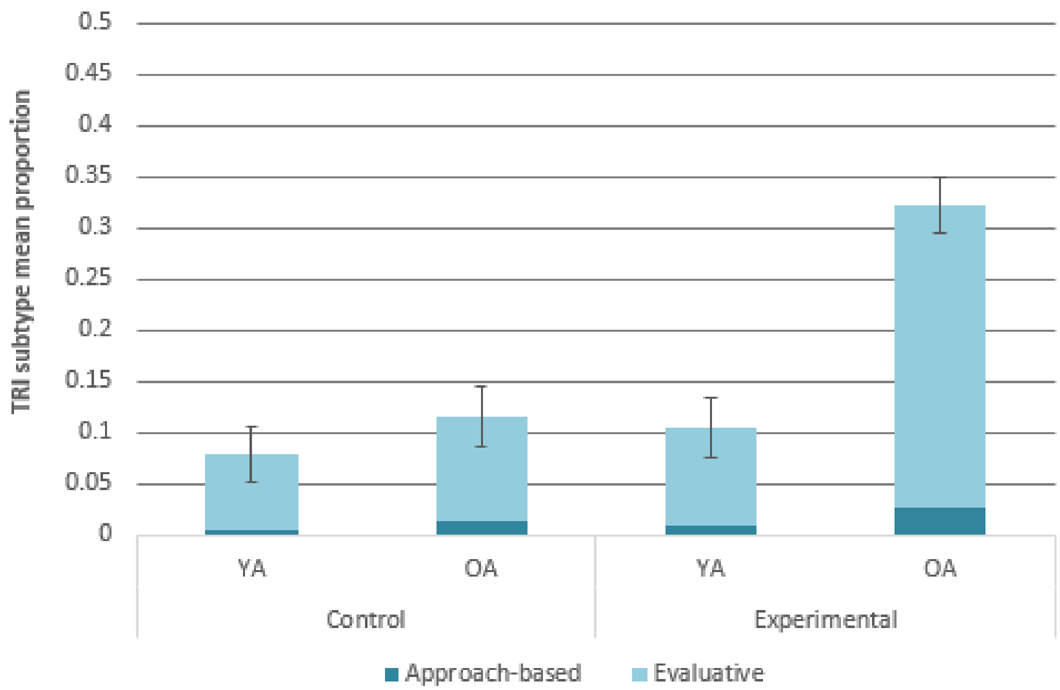


Figure 9. TRI for Study 5.

The different subtypes of approach-based TRI (Figure 9) for younger and older adults are shown below. While we expected main effects of age and condition on overall proportion of TRI, we did not make specific predictions regarding the effect of age and condition on approach-based TRI or the different subtypes of approach-based TRI. There was an effect of condition on proportion of approach-based TRI, F (1, 215) = 6.231, p = .013, $\eta p2$ = .010. JOL + probe participants (M = .019, SE = .003) reported more approach-based TRI than probe only participants (M = .007, SE = .002), (t (210) = 2.579, p =.011, d = .453). Having participants respond to JOL screens resulted in them experiencing more off-task thoughts about task approach and strategy.

There was no effect of age on approach-based TRI (F (1, 215) = .774, p = .461, $\eta p2$ = .002) and no age x condition interaction (F (1, 215) = .258, p = .773, $\eta p2$ = .001).

However, there was an effect of approach-based TRI *subtype* on proportion of approach-based TRI reported, (*F* (1, 215) = 20.610, $p < .001$, $\eta p2$ = .061). Participants reported more TRI about new strategies (*M* = .130, *SE* = .021) they could use to complete the task than TRI about the effectiveness (*M* = .016, *SE* = .007) of their current strategies, *t* (210) = 5.150, $p < .001$, *d* = .701. Likewise, participants reported more TRI about new strategies (*M* = .130, *SE* = .016) they could use to complete the task than TRI about the ineffectiveness (*M* = .032, *SE* = .011) of their current strategies, *t* (210) = 4.134, $p < .001$, *d* = .688. There was no difference in TRI about effective (*M* = .014, *SE* = .006) and ineffective strategies (*M* = .030, *SE* = .010), (*t* (210) = 1.372, *p* = .172, *d* = .082).

While there were no significant age x TRI subtype (*F* (1, 215) = .774, *p* = .461, $\eta p2$ = .002) or age x condition x TRI subtype interactions (*F* (1, 215) = .258, *p* = .773, $\eta p2$ = .001), there was a significant condition x TRI subtype interaction, *F* (1, 215) = 6.342, *p* = .002, $\eta p2$ = .020. Probe only participants (*M* = .069, *SE* = .023) reported fewer thoughts about adopting new task strategies than JOL + probe participants (*M* = .185, *SE* = .036), *t* (210) = 2.682, *p* =.008, *d* = .370. Responding to JOL screens seems to prompt participants to think about new strategies they could use to successfully complete the paired associates task.

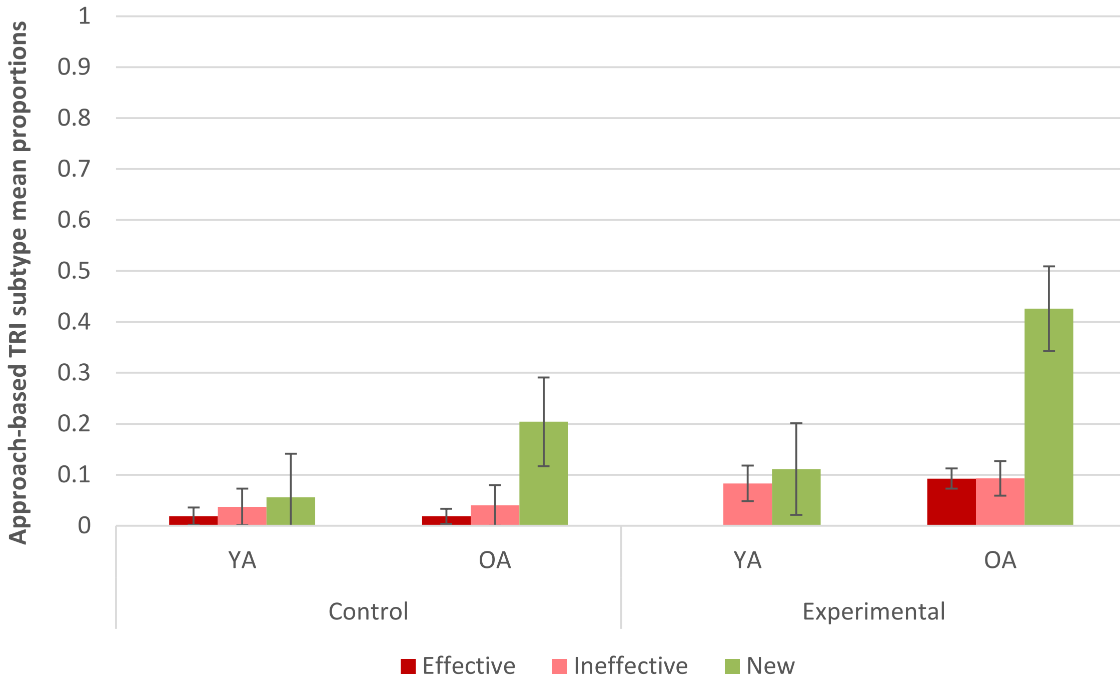


Figure 10. Approach-Based TRI for Study 5.

The different subtypes of evaluative TRI (Figure 10) for younger and older adults are shown below. While we expected main effects of age and condition on overall proportion of TRI, we did not make specific predictions regarding the effect of age and condition on evaluative TRI. There was an effect of age on proportion of evaluative TRI, F (1, 215) = 7.438, p = .007, $\eta p2$ = .009. The proportion of evaluative TRI was significantly higher for older adults (M = .196, SE = .026) than it was for younger adults (M = .088, SE = .013), t (210) = 3.715, p < .001, d = .506). There was also an effect of condition on proportion of evaluative TRI, F (1, 215) = 5.351, p = .021, $\eta p2$ = .006. JOL + probe participants (M = .199, SE = .026) reported more evaluative TRI than probe only participants (M = .085, SE = .013), (t (210) = 3.922, p < .001, d = .534). The age x condition interaction was not significant (F (1, 215) = .746, p = .389, $\eta p2$ = .004).

There was also an effect of evaluative TRI *subtype* on proportion of evaluative TRI reported, (F (1, 215) = 31.098, p < .001, $\eta p2$ = .099). Participants reported more TRI about task ability (M = .165, SE = .019) than TRI about their current task performance (M = .075, SE = .014), t (210) = 3.813, p < .001, d = .520, or about their overall task performance (M = .059, SE = .012), t (210) = 4.717, p < .001, d = .643. Participants also reported more TRI about task difficulty (M = .287, SE = .026) than TRI about current task performance (M = .156, SE = .015), t (210) = 7.063, p < .001, d = 1.007, overall task performance (M = .125, SE = .012), t (210) = 7.962, p < .001, d = 1.084, or task ability (M = .165, SE = .019), t (214) = 3.789, p < .001, d = .516. The difference between TRI about overall and current task performance was not significant, t (210) = .833, p = .406, d = .546.

While there were no significant age x condition (F (1, 215) = 1.487, p = .223, $\eta p2$ = .002) or age x condition x TRI subtype interactions (F (1, 215) = .533, p = .660, $\eta p2$ = .002), there was a significant age x TRI subtype interaction, F (1, 215) = 3.019, p = .029, $\eta p2$ = .011. Older adults (M = .212, SE = .027) reported more TRI about task ability than younger adults (M = .118, SE = .026), t (210) = 2.508, p = .013, d = .341. It is possible that older adults, who could be coming into the lab with concerns about cognitive decline, had more worry-laden thoughts about their cognitive ability than younger adults (Hertzog & Hultsch, 2000).

Older adults (M = .344, SE = .038) also reported more TRI about task difficulty than younger adults (M = .230, SE = .034), t (210) = 2.236, p = .026, d = .304. Again, concerns about cognitive decline in older adults could have resulted in them reflecting more on the difficulty of the task compared to younger adults. Older and younger adults did not differ in their reported TRI about current performance (*F* (1, 215) = .013, *p* = .909, *ηp2* = .000) or overall performance (*F* (1, 215) = .083, *p* = .774, *ηp2* = .000).

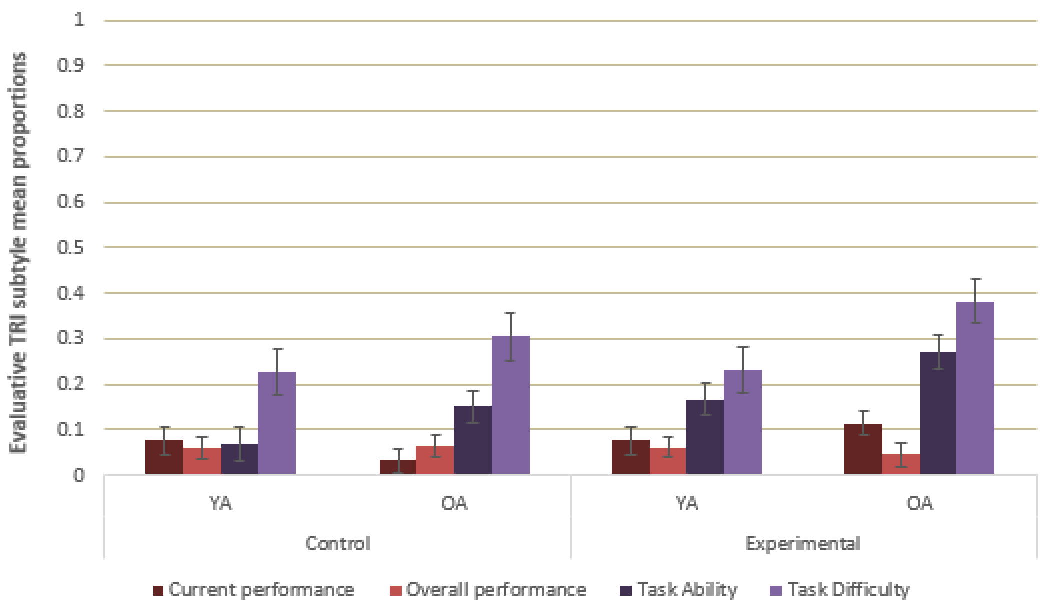


Figure 11. Evaluative TRI for Study 5.

Interestingly, the proportion of probe-caught “other TRI” was significantly higher for younger adults (*M* = .091, *SE* = .018) than it was for older adults (*M* = .022, *SE* = .019), *F* (1, 208) = .495, *p* = .002, *ηp2* = .034). Again, off-task thoughts within the “other TRI” category generally encompassed thoughts regarding how long the task is and how much time the task is taking. There was no main effect of condition on “other TRI”, *F* (1, 215) = .467, *p* = .495, *ηp2* = .002). Likewise, there was no age x condition interaction, *F* (1, 208) = .001, *p* = .978, *ηp2* = .000).

We conducted post-hoc analyses to examine the effect of block on approach-based and evaluative TRI. Descriptive statistics for approach-based and evaluative TRI broken down by age group, condition, and task block are presented in Appendix R. In addition to the finding that both younger and older adults report more TRI during the first task block we also found that participants reported more approach-based TRI during Block 1 (*M* = .010, *SE* = .002) compared to Block 2 (*M* = .003, *SE* = .001), F (1, 212) = 12.944, $p < .001$, $\eta p2$ = .058. Participants also reported more evaluative TRI during Block 1 (*M* = .083, *SE* = .008) compared to Block 2 (*M* = .059, *SE* = .007), F (1, 212) = 19.611, $p = .001$, $\eta p2$ = .085. For both approach-based and evaluative TRI there were no block x age, block x condition, or block x age x condition interactions (all *p*-values > .05).

Descriptive statistics for approach-based and evaluative TRI broken down by age group, condition, and paired associates item relatedness are presented in Appendix S. There was a main effect of item relatedness of proportion of approach-based TRI, F (1, 212) = 9.376, $p = .002$, $\eta p2$ = .042. Pairwise comparisons showed that participants reported significantly more approach-based TRI on trials where they had just studied an item pair that was unrelated (*M* = .008, *SE* = .001) compared to trials where the word pairs preceding the probe were moderately related (*M* = .001, *SE* = .001) or strongly related (*M* = .003, *SE* = .001) (*p*-values < .05). There was no difference between moderately related and strongly related words in terms of approach-based TRI (*p*-value > .05).

There was also main effect of item relatedness of proportion of evaluative TRI, *F* (1, 212) = 16.846, $p < .001$, $\eta p2 = .075$. As with approach-based TRI, pairwise comparisons showed that participants reported participants reported significantly more evaluative TRI on trials where they had just studied an unrelated pair (*M* = .060, *SE* = .005) compared to probes following moderately related (*M* = .039, *SE* = .004) and unrelated (*M* = .060, *SE* = .005) pairs. There was no difference in evaluative-based TRI for moderately related and strongly related trials ($p > .05$). These differences in overall TRI, approach-based TRI, and evaluative TRI as a function of item relatedness are perhaps unsurprisingly, as both younger and older adults have been found to report more TRI on difficult tasks compared to easier tasks (McVay et al., 2013). Seeing unrelated word pairs, which are more difficult to remember, resulted in participants engaging in more monitoring.

Recall Performance

To examine the effects of condition and age on paired associates recall performance, we used a 2 (age: young, old) x 2 (condition: probe only, JOL + probe) factorial ANOVA with age and condition as between-subject variables. In keeping with previous work (Hertzog, Kidder, Powell-Moman, Dunlosky, 2002; Hertzog, Sinclair, & & Dunlosky, 2010) we expected that, in both conditions, younger adults would correctly recall more response words than older adults. For recall performance, we did not necessarily expect a main effect of condition or an age x condition interaction.

Although we expected that participants within our JOL condition would engage in more monitoring, increases in monitoring might not result in immediate increases in recall performance.

For paired associates recall performance, older adults (*M* = .449, *SE* = .019) had significantly lower paired associates recall accuracy than younger adults (*M* = .537, *SE* = .018), F (1, 215) = 10.376, p = .001, $\eta p2$ = .045), as expected. Contrary to our initial hypothesis, there was a main effect of condition on paired associates recall (F (1, 215) = 9.883, p = .002, $\eta p2$ = .044), with participants in the JOL + probe condition (*M* = .535, *SE* = .019) had significantly higher paired associates recall accuracy than probe only participants (*M* = .450, *SE* = .018). This finding is interesting, as other work has found that having participants make metacognitive judgements does not result in reactivity in terms of paired associates recall performance when the word list used consists of a mixture between related and unrelated words (see Double, Birney, & Walker, 2017 for a review of this literature).

While the age x condition interaction did not reach statistical significance (F (1, 215) = 3.541, p = .061, $\eta p2$ = .015), it appears that making JOLs resulted in greater increases in recall performance for younger adults compared to older adults. Younger adults in the JOL + probe condition (*M* = .580, *SE* = .028) had numerically higher recall performance than younger adults in the probe only condition (*M* = .494, *SE* = .026). Older adults in the JOL + probe condition (*M* = .495, *SE* = .028) and older adults in the probe only condition (*M* = .490, *SE* = .029) had very similar recall performance.

To examine the reliability of our paired associates recall measure across Block 1 and Block 2 of the paired associates task we calculated split-half reliability for recall. The Spearman-Brown split-half reliability coefficient for paired associates recall performance across the two task blocks was .79, indicating *adequate* reliability (Hulin, Netemeyer, & Cudeck, 2001). As participants completed the Block 1 recall task participants may have received implicit feedback regarding the effectiveness of ineffectiveness of their task strategies which could have resulted in them using different, more effective strategies during Block 2 compared to Block 1. Differences in task strategy across block could result in differences in task performance. Descriptive statistics for paired associates recall performance broken down by age group, condition, and task block are presented in Appendix R. We did not find an effect of block on PA recall performance, F (1, 212) = 501.293, p = .260, $\eta p2$ = .006. Participants had comparable performance on Block 1 (M = .457, SE = .014) and Block 2 (M = .447, SE = .015) of the task. Likewise, the block x age and block x age x condition interactions were not significant (p-values > .05). Completing a recall task before studying the Block 2 word pairs did not result in improved Block 2 performance.

Descriptive statistics for paired associates recall performance broken down by age group, condition, and paired associates item relatedness are presented in Appendix S. There was a main effect of item relatedness on recall performance, F (1, 212) = 1.278, p = .260, $\eta p2$ = .703. Unsurprisingly, pairwise comparisons showed that participants had significantly lower recall for unrelated items (M = .280, SE = .015) compared to moderately related (M = .489, SE = .014) and strongly related items

(*M* = .582, *SE* = .014) (*p*-values < .05) and that participants also had significantly lower recall for moderately related items (*M* = .489, *SE* = .014) compared to strongly related items (*M* = .582, *SE* = .014) (*p*-value < .05). The age x relatedness, condition x relatedness, and age x condition x relatedness interactions were not significant (all *p*-values > .05).

JOLs

For our JOL + probe condition, we examined age differences in mean levels of JOLs, absolute accuracy of JOLs, and relative accuracy of JOLs. A summary of the descriptive statistics for these variables are presented Table 1 below. Younger adults (*M* = 54.063, *SE* = 2.387) and older adults (*M* = 17.637, *SE* = 2.400) had similar mean JOLs, t (106) = .387, p = .700, d = 113. This is inconsistent to past work that has found that younger adults report higher mean JOLs than older adults (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002). Post-hoc analyses revealed that participants had higher mean JOLs for Block 1 (*M* = 54.789, *SE* = 1.877) than Block 2 (*M* = 51.299, *SE* = 1.982), F (1, 106) = 13.431, p < .001, $\eta p2$ = 112). Participants appeared to have gotten more confident in their ability to correctly recall the words as the task went on. We did not find a block x age interaction for mean JOLs (*p*-value > .05).

There was a main effect of relatedness on participants' mean JOLs, F (1, 107) = 125.675, p < .001, $\eta p2$ = .705. Consistent with past work (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002), pairwise comparisons showed that participants gave higher JOLs to strongly related word pairs (*M* = .656, *SE* = .021) compared to moderately related (*M* = .598, *SE* = .022) and unrelated (*M* = .319, *SE* = .020) word pairs.

Participants also had significantly higher mean JOLs for moderately related items (*M* = .598, *SE* = .022) compared to unrelated items (*M* = .319, *SE* = .020). We did *not* find an age x relatedness interaction for mean JOLs (*p*-value > .05). While both younger and older adults gave the highest JOLs to strongly related word pairs and the lowest JOLs to the unrelated word pairs (see Appendix S), this effect was not greater in older adults. Therefore, we did *not* replicate past work (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002) showing that sensitivity to item relatedness was greater for older adults than younger adults.

The simple difference between a participants' mean JOL and mean recall measures their average overconfidence and underconfidence. We expected a possible effect of age on the difference scores, with older adults having mean JOLs closer to their actual mean recall levels (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002). While older adults did have mean JOLs closer to their actual mean recall, younger (*M* = .064, *SE* = .025) and older (*M* = -.014, *SE* = .025) adults were not significantly different. Both younger and older adults were well-calibrated in terms of their JOLs, and had little discrepancy between their mean JOLs and mean recall. There was no difference between younger and older adults for the simple difference between mean JOL and mean recall, *t* (106) = 1.935, *p* = .056, *d* = .425.

Simple difference scores were similar in Block 1 (*M* = -.036, *SE* = -.033) and Block 2 (*M* = .171, *SE* = .029), *F* (1, 106) = .039, *p* = .845, *ηp2* = .002). Block was not found to interact with age (*p*-value > .05). There was a main effect if item relatedness on participants' simple difference scores, *F* (107) = 8.124, *p* = .001, *ηp2* = .134. Participants had smaller difference scores for strongly related items (*M* = -.003, *SE* = .024) compared to moderately related items (*M* = -.054, *SE* = .023), indicating that participants were more accurate in monitoring their performance when the word pairs were more strongly related. The age x relatedness interaction was not reliable (*p*-values > .05), in keeping with past work that has looked at the effect of item relatedness on JOL simple difference scores.

The absolute difference or discrepancy between mean JOL and recall (regardless of whether the JOLs underestimate or overestimate recall) is a measure of absolute metacognitive accuracy used within the metacognition literature (Devolder et al., 1990). We expected smaller absolute difference scores than younger adults (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002) compared to older adults. However, younger (*M* = .162, *SE* = .015) and older adults (*M* = .199, *SE* = .016) had comparable absolute difference scores, t (102) = 1.671, p = .098, d = .324. Participants had similar absolute difference scores for Block 1 (*M* = .172, *SE* = .012) and Block 2 (*M* = .189, *SE* = .017), F (1, 106) = .229, p = .633, $\eta p2$ = .002). The age x block interaction was not reliable (*p*-value > .05). In keeping with past work, the main effect of relatedness and the age x relatedness interaction on absolute difference scores were not significant (both p-values > .05).

Table 1

Monitoring Indices Descriptive Statistics

	YA Probe Only		***OA Probe Only***		***p-value***
N	*53*		*53*		
	M	*SE*	*M*	*SE*	
Mean JOLs	54.063	2.387	52.083	2.400	.049
JOL Simple Diff	.064	.024	-.014	.025	.056
JOL Absolute Diff	.162	.015	.199	.016	.098
Gamma	.273	.022	.241	.023	322

Note. Only experimental participants were required to make JOLs in addition to responding to thought content probes.

We were also able to examine the effects of age on JOL resolution within our JOL + probe condition. Because it is one of the most widely used measures of metacognitive accuracy within the memory literature, relative accuracy of JOLs was operationalized in this study as the gamma correlations (Goodman & Kruskal, 1954; Nelson, 1984) between JOLs and recall test performance. Computing gamma correlations will allow us to compare the results of this study to other studies within the metacognition literature. Because extreme marginal values of JOLs or recall can distort measures of metacognitive accuracy, participants who recalled less than 5% correct or greater than 95% correct of the items were excluded from our analyses, as such, two younger adults were excluded from the following analyses. No participants needed to be dropped due to having insufficient variability in their JOL responses. Given previous research, we did not expect an effect of age on JOL resolution (Hertzog, Sinclair, & Dunlosky, 2010). As expected, the gamma correlations for younger adults within the JOL + probe condition (*M* = .273, *SE* = .022) and older adults within the JOL + probe condition (*M* = .241, *SE* = .023) did not differ (*t* (102) = .995, *p* = .322, *d* = .193. Older adults were as accurate as younger adults in estimating their likelihood of recalling the paired associates task stimuli.

Participants had similar gamma coefficients for Block 1 (*M* = .261, *SE* = .021) and Block 2 (*M* = .253, *SE* = .018), *F* (1, 567) = .163, *p* = .687, *ηp2* = .002). The block x age interaction was not significant (*p*-value > .05). The main effect of relatedness and age x relatedness interaction were both non-significant (both p-values > .05). JOL resolution did not change throughout the task and did not differ according to item difficulty. This finding is consistent with other work that has shown that relatedness does not effect relative accuracy in younger and older adults (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002).

Within the JOL + probe condition, we conducted exploratory analyses to measure the association between monitoring frequency and monitoring accuracy by measuring the correlation between proportion of TRI and gamma correlations. An a priori power analysis in G*Power (Faul, Erdfelder, Buchner, & Lang, 2007) suggested that a total sample of 106 participants (53 participants in each age group within the JOL + probe condition) should be sufficient to achieve 89% power to detect a medium effect size when looking at the correlation between TRI frequency and JOL resolution. This sample size calculation was conducted using the traditional .05 criterion of statistical significance. Correlations between our various monitoring indices (gamma, mean JOL, and absolute difference between recall performance and JOLs), TRI, use of effective recall strategies, and PA task recall performance are presented in Table 2 below. Correlations between the monitoring indices, TRI, use of effective recall strategies, and

PA task recall performance within the younger adult sample and the older adult sample are presented in Table 3 and Table 4, respectively.

We expected that, both overall and within each age group within the JOL + probe condition, proportion of TRI would be positively associated with the gamma correlations, indicating that those who monitor performance more frequently will also monitor with greater accuracy. We also expected a negative association between proportion of TRI and absolute difference scores between mean JOLs and mean recall performance, indicating that those who monitor performance more frequently also are better able to predict their eventual recall performance. Contrary to our predictions, gamma correlations was not correlated with TRI[1] (r = -.064, p = .508), JOL difference scores (r = -.011, p = .914), or absolute differences scores (r = .011, p = .914). Within the younger adult sample and within the older adult sample, gamma correlations, JOL difference scores, and JOL absolute difference scores likewise were not associated with TRI (all p-values > .05). Our measures of metacognitive monitoring accuracy were not associated with evaluative TRI and approach-based TRI (all p-values > .05). Individuals who reported more frequent monitoring did not monitor more accurately than those who reported less frequent monitoring.

Table 2

Overall Correlations Between Monitoring Indices, TRI, PA Recall, and PA Strategy Use

	PA Recall	**Overall TRI**	**Approach-based TRI**	**Evaluative TRI**	**JOL Abs Diff**	**Gamma**	**Effective Strat**
PA Recall	1	-.064	.005	-.068	.041	.226*	.306**
Overall TRI		1	.294**	.988**	.011	-.064	.076
Approach-Based TRI			1	.142*	.069	-.088	-.083
Evaluative TRI				1	070	-.052	.059
JOL Abs Diff					1	.003	.137
Gamma						1	-.056
Effective Strategy							1

Note. Effective Strat refers to use of use of effective mediator-based strategies during Block 2 of the PA task; * = $p < .05$ and ** = $p < .01$.

Table 3

Correlations Between Monitoring Indices, TRI, PA Recall, and Strategy Use for YAs

	PA Recall	Overall TRI	Approach-based TRI	Evaluative TRI	JOL Abs Diff	Gamma	Effective Strat
PA Recall	1	.008	.027	.004	.042	.307*	.378**
Overall TRI		1	.399**	.987**	.163	-.130	.061
Approach-Based TRI			1	.248*	.181	-.161	-.032
Evaluative TRI				1	.141	-.108	.070
JOL Abs Diff					1	.198	.208
Gamma						1	.054
Effective Strategy							1

Note. Effective Strat refers to use of use of effective mediator-based strategies during Block 2 of the PA task; * = $p < .05$ and ** = $p < .01$.

Table 4

Correlations Between Monitoring Indices, TRI, PA Recall, and PA Strategy Use for OAs

	PA Recall	**Overall TRI**	**Approach-based TRI**	**Evaluative TRI**	**JOL Abs Diff**	**Gamma**	**Effective Strat**
PA Recall	1	-.078	.015	-.083	.113	.120	.217*
Overall TRI		1	.241*	.986**	.022	.011	-.330**
Approach-Based TRI			1	.084	-.070	-.038	.001
Evaluative TRI				1	.036	.018	.120
JOL Abs Diff					1	-.160	.081
Gamma						1	-.165
Effective Strat							1

Note. Effective Strat refers to use of use of effective mediator-based strategies during Block 2 of the PA task; * = $p < .05$ and ** = $p < .01$.

Strategy Use

As mentioned earlier, participants in both conditions completed strategy reports following the Block 2 recall portion of the paired associates task. We analyzed the effect of strategy use by aggregating the likelihood of participants using an effective mediator-based strategy (imagery use and sentence generation; see Hertzog et al., 2009) into a single variable and calculating the proportion of recall trials where an effective strategy was endorsed. Replicating previous work, we expected a significant effect of age on strategy use, with older adults using effective strategies less often than younger adults.

We found that younger (M = .578, SE = .023) and older adults (M = .544, SE = .021) reported similar use of effective strategies, (F (1, 215) = 1.385, p = .241, $\eta p2$ = .006). This lack of age difference in use of effective strategies is surprising, but might account for why older adults performed nearly as well on the recall portion of the paired associates task as younger adults. There were no condition differences (F (1, 215) = .086, p = .770, $\eta p2$ = .000) in use of effective learning strategies, and there was not an age x condition interaction (F (1, 215) = .368, p = .545, $\eta p2$ = .002).

We initially believed it possible that individuals who engage in more metacognitive monitoring and who report more TRI would also generate more effective strategies during the paired associates tasks than individuals who engage in less frequent metacognitive monitoring. We found that, overall, frequency of monitoring (TRI) and efficient strategy use on Block 2 of the paired associates task were not significantly correlated with each other (r = .076, p = .275). Within both the younger adult (r = .059, p = .182) and older adult samples (r = .061, p = .546), TRI and efficient strategy use remained uncorrelated. Likewise, frequency of approach-based TRI was not significantly correlated with efficient strategy use (r = -.083, p = .348).

Post-hoc analyses revealed a main effect of item relatedness on use of effective memory strategies, F (1, 212) = 110.780, p < .001, $\eta p2$ = .343). Pairwise comparisons showed that participants endorsed using effective strategies more often when the to-be-remembered items were strongly related (M = .586, SE = .008) compared to when they were moderately related (M = .563, SE = .007) or unrelated (M = .473, SE = .010). This likely explains why participants had the highest recall performance for strongly related items. We did not find age x relatedness, condition x relatedness, or age x condition x relatedness interactions for use of effective strategies (p-value > .05).

Post-Task Questionnaires

As mentioned, we administered a variety of post-task questionnaires following completion of the paired associates task, including several single-item Likert scale questions assessing perceived stress, fatigue, distractibility, motivation, interest, and difficulty during the paired associates task. Means and standard errors for post-task questionnaires are presented in Table 5 below. adults (M = 3.300, SE = .115) reported feeling more fatigue during the paired associates task than did older adults (M = 2.375, SE = .117), (F (1, 215) = 31.043, p < .001, $\eta p2$ = .129). There was no effect of condition on self-reported fatigue, and no age x condition interaction (all p-values > .05).

There was also an effect of age on task interest, with older adults (M = 2.907, SE = .107) reporting that they found the paired associates task to be more interesting than did younger adults (M = 1.870, SE = .104), (F (1, 215) = 47.258, p < .001, $\eta p2$ = .184). This is typical, with older adults reporting increased task interest relative to younger adults on a variety of cognitive tasks (Jackson & Balota, 2012; Krawietz, Tamplin, & Radvansky, 2012; Frank et al., 2015; Jordano & Touron, 2017a) in There was no main effect of condition on self-reported task interest, and no age x condition interaction (*p*-values > .05).

Finally, there was an effect of age on self-rated distractibility during the paired associates task, with younger adults (*M* = 3.115, *SE* = .107) reporting that they got distracted more easily than did older adults (*M* = 1.804, *SE* = .112), (*F* (1, 215) = 69.537, *p* < .001, $\eta p2$ = .184). There was also main effect of condition on self-reported distractibility, with those in the probe only condition (*M* = 2.634, *SE* = .109) reporting more distractibility than those in the JOL + probe condition (*M* = 2.634, *SE* = .111), (*F* (1, 215) = 4.915, *p* = .028, $\eta p2$ = .249). The age x condition interaction was also significant, (*F* (1, 215) = 5.347, *p* = .022). Younger adults in the probe only condition (*M* = 3.472, *SE* = .165) had higher perceived distractibility than younger adults that had to respond to JOL screens (*M* = 2.759, *SE* = .201), *t* (104) = 2.731, *p* = .007, *d* = .373. Older adults in the probe only condition (*M* = 1.796, *SE* = .133) did not have higher perceived distractibility than older adults that had to respond to JOL screens (*M* = 1.811, *SE* = .114), *t* (104) = .085, *p* = .932, *d* = .012.

Finally, we administered a modified version of the Cognitive Failures Questionnaire (CFQ) as a measure of everyday monitoring failures. Past studies using the CFQ have found evidence of underreporting of memory errors in older adults, with older adults reporting a similar overall number of everyday memory failures as younger adults

(Kramer, Humphrey, Larish, Logan, & Strayer. 1994; Reese & Cherry, 2006) or older adults reporting fewer everyday memory errors than younger adults (Mecacci & Righi, 2006). In the current study, there were age differences in reported everyday memory monitoring failures, with older adults (*M* = 71.252, *SE* = 1.219) obtaining higher scores on the CFQ compared to younger adults (*M* = 53.760, *SE* = 1.278), (F (1, 215) = 104.466, p = < .001, $\eta p2$ = .332). It has been suggested that age-equivalence and age-related declines in CFQ scores could reflect decreased ability of older adults to remember their own cognitive failures (de Winter, Dodou, & Hancock, 2015). The fact that older adults in the current study reported more everyday cognitive failures may indicate that our older sample was able to accurately reflect upon their everyday memory performance and thus report more cognitive failures. There was no effect of age, condition, or an age x condition interaction on DSSQ scores (all *p*-values > .05).

Table 5

Post-Task Questionnaire Data

	YA Probe Only		*YA JOL + Probe*		*OA Probe Only*		*OA JOL + Probe*	
N	*53*		*53*		*53*		*53*	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
DSSQ TRI	22.96	.64	21.65	.63	20.72	.63	22.08	.64
CFQ	51.46	1.21	56.06	1.24	70.28	1.25	72.23	1.30
PTQ Task Difficulty	2.79	.13	3.17	.12	2.96	.13	3.09	.12
PTQ Fatigue	3.41	.17	2.22	.16	3.18	.17	2.53	.17
PTQ Instruction	2.45	.15	2.13	.15	2.22	.15	1.90	.14
PTQ Motivation	2.21	.16	2.48	.17	2.70	.15	2.81	.16
PTQ Stress	2.28	.15	2.67	.16	2.02	.16	2.19	.17
PTQ Interest	1.81	.15	1.92	.16	3.00	.17	2.82	.14
PTQ Effort	1.96	.13	1.85	.13	2.57	.13	2.46	.14
PTQ Satisfy	2.81	.14	2.69	.13	3.17	.13	3.06	.15
PTQ Distractibility	3.47	.16	1.80	.15	2.76	.15	1.81	.16

Note. Scores on the modified PTQ could range from 8 to 40, with higher scores representing more everyday TRI. Scores on the Cognitive Failures Question (CFQ) could range from 0 to 100 with higher scores representing more instances of everyday memory failures. All responses to the post-task questions above were made on a 1 to 5 Likert scale. For PTQ Difficulty, 1 = Not at all difficult and 5 = Very difficult. For PTQ Fatigue, 1 = Not at all fatiguing and 5 = Very fatiguing. For PTQ Instruction (“To what extent do you feel like you understood the task instructions?”), PTQ Motivation, PTQ Stress, PTQ Fatigue, PTQ Interest, PTQ Satisfy (“How much are you satisfied in your task performance?”), and PTQ Desirability, 1 = Not at all and 5 = Very much. For PTQ Effort (“How much effort did you put into this study?”, 1 = Very little effort and 5 = A lot of effort.

CHAPTER X

STUDY 5 DISCUSSION

Research that provides insight into the frequency of spontaneous metacognitive monitoring is sparse. 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 unknown exactly how frequently people engage in thoughts regarding their task performance on both laboratory and everyday cognitive tasks. Furthermore, it is possible that having participants make explicit metacognitive judgments such as JOLs leads to reactivity. The proposed study investigated the frequency of metacognitive monitoring in younger and older adults during a commonly-used memory task and investigated the effect that making JOLs has on participants' monitoring frequency and their eventual task performance and task strategies.

As mentioned within the Introduction, studying metacognition typically requires participants to make *explicit* predictions or judgments about their cognitive abilities. While such judgments are necessary to properly study metacognitive accuracy, it is possible that individuals participating in metacognition studies engage in monitoring and other types of metacognitive thinking to a greater degree because they are being asked by the experimenter or task instructions to make performance judgments. Within the proposed study, a group of participants who were not asked to provide explicit metacognitive judgments (our probe only condition) were compared to a group that makes traditional metacognitive judgments (our JOL + probe condition) in terms of monitoring accuracy.

We found that, at least for older adults, requiring participants to make explicit judgements of learning results in increased monitoring. While we did not find that increased monitoring by older adults who made JOLs resulted in changes in strategy use and eventual recall performance, it is possible that requiring older adults to make JOLs could lead to changes in study strategies and performance in other types of cognitive tasks. Additionally, participants in the JOL + probe condition had significantly higher paired associates recall accuracy than participants that were not required to make JOLs, despite not reporting increased use of effective study strategies. While the age x condition interaction was only trending, the difference in recall performance between participants who only responded to thought probes and participants who responded to thought probes and provided JOLs was greater in the younger adult sample relative to the older adult sample. As such, including instructions or task demands aimed at increasing one's propensity to engage in monitoring may be a way to boost task performance in younger adults who may not otherwise engage in monitoring.

The current study is also the first known study to examine the impact of making metacognitive judgements on paired associate performance in older adults. Providing

JOLs resulted in participants reporting more thinking about both task strategy and their task performance and the effect of providing JOLs on the frequency of metacognitive thinking was greater for older adults than it was for younger adults. Older adults are presumably coming into the testing environment with more concerns regarding cognitive performance and cognitive decline than younger adults (Hertzog & Hultsch, 2002; Jordano & Touron, 2017a), particularly when they know they are about to complete a task with a memory component. Requiring older adult participants to provide metacognitive judgements during a memory task may further increase worries regarding memory ability and memory performance. While we did not assess the *emotional valence* of participants' metacognitive thoughts in the current study, further work can examine whether requiring metacognitive judgements in older adults specifically increases negative appraisals of one's task ability and performance.

One might expect that participants would report different types of off-task thoughts according to the type of task they are performing. Previous studies of mind-wandering have assessed mind-wandering during reading tasks (Jackson & Balota, 2012; McVay & Kane, 2012; Kraweitz, Tamplin, & Radvansky, 2012; Unsworth & McMillan, 2013; Frank et al., 2015), sustained attention tasks (McVay & Kane, 2009; Jackson & Balota, 2012), and span tasks (Jordano & Touron, 2017a, Jordano & Touron, 2017b). The current study was the first study to assess mind-wandering during an associative memory task. We replicated the common finding that older adults are on task more than younger adults and experience more TRI than do younger adults (McVay et al., 2013; Zavagnin et al., 2014; Frank et al., 2015; Jordano & Touron, 2017a), and that the TRI reported by younger and older adults encompasses metacognitive thoughts regarding the effectiveness of task strategy and one's task performance (Jordano & Touron, 2017a). Likewise, we replicated the typical finding that younger adults experience more TUTs relative to older adults (McVay & Kane, 2009; Jackson & Balota, 2012; McVay & Kane, 2012; McVay et al., 2013; Unsworth & McMillan, 2013; Jordano & Touron, 2017a). Given that TRI seems to be analogous to metacognitive monitoring, the results of the study suggests that older adults engage in more spontaneous monitoring than younger adults during associative memory task, although they do not experience more spontaneous off-task thoughts that are unrelated to the current task. Both the frequency of metacognitive monitoring and the *specific content* of individuals' TRI experiences can vary according to characteristics of the ongoing task. Additional work can be done to determine how aspects of the task affect participants' propensity to monitoring.

Focused studies examining spontaneous monitoring may also be used to further investigate the relationship between monitoring frequency and monitoring accuracy. In the current study, we replicated previous work demonstrating age invariance in monitoring resolution (Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002). We initially expected that individuals who more frequently monitor and who therefore have more practice monitoring would also monitor more accurately than those who monitor less frequently, and that age-related increases in monitoring may explain the aforementioned age invariance in monitoring ability. By having participants in one of our conditions complete both thought content probes and metacognitive judgments during the same task, we were able begin to address this question. Contrary to our hypotheses, both younger and older adults who reported more monitoring during the paired associates task did not have greater monitoring accuracy than those who reported less monitoring.

It is possible that participants that engage in monitoring frequently do not build up monitoring expertise and that factors other than monitoring frequency account for older adults' spared metacognitive monitoring. Both younger and older adults can effectively use similar encoding cues when making JOLs (Hertzog, Dunlosky, Robinson, & Kidder, 2003; Robinson, Hertzog, & Dunlosky, 2006). Because older adults are aware of age-related declines in memory ability, they might strategically modify their attention and goal-directed processing in a way that allows them to attend to cues and make accurate metacognitive judgments (Castel, McGillivray, & Friedman, 2012; Hertzog & Dunlosky, 2011). The theory that older adults selectively regulate their limited cognitive resources to better attend to task-related cues is in keeping with theories of cognitive aging that are used to explain other examples of age-invariance in cognitive performance (theory of selective optimization with compensation; Baltes & Baltes, 1990). Additional studies are needed to determine why older adults can attend to task-related cues as well as younger adults and make metacognitive judgments that are as accurate as those of younger adults.

The current study has both strengths and limitations. As mentioned, we extended work on mind-wandering to a commonly used associative memory task. The results of this study provide insight into how frequently individuals mind-wander and what individuals mind-wander about during a learning and memory task that allows participants to adopt a variety of strategies to complete the task. This study also extended work examining reactivity associated with having participants make metacognitive judgements during an ongoing task. Finally, within the current study we begin to address questions regarding metacognitive monitoring frequency and the effect of individual differences in monitoring frequency on metacognitive accuracy, task strategy, and task performance.

The current study is not without limitations and there are numerous avenues for future research. One surprising finding within the current study is that both younger and older adults had lower paired associates recall performance than has been found in other studies using similar task stimuli (Hertzog, Sinclair, & Dunlosky, 2010; Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002), despite both younger and older adults reporting high use of effective mnemonic strategies. The paired associates task was longer than ones used in these previous studies, but task performance was not lower on Block 2 of the task relative to Block 1. Therefore, this lower recall performance does not seem to be due to factors such as fatigue associates with task length. We are unsure of why younger and older adults had paired associate recall performance that was lower than what has been found in other studies, but one possible explanation could be that having participants report to online mind-wandering probes altered paired associates processing, decreasing task performance. Further work can be done investigating how responding to mind-wandering thought probes alters task performance.

An additional limitation of the current study is that we only assessed metacognitive monitoring frequency during encoding. As mentioned within the Introduction, individuals are believed to engage in monitoring during both encoding and retrieval and monitoring accuracy during retrieval can be assessed by having participants make FOK judgements. It is possible that individuals monitor to a greater or lesser extent, or engage in mind-wandering about different things, during the retrieval stages of an ongoing task. Additional studies can be completed investigating how frequency of monitoring changes across different stages of a memory task and whether or not frequency of monitoring is associated with accuracy of other metacognitive judgements, such as FOKs and confidence judgements. The frequency of metacognitive monitoring in younger and older adults and its relationship to metacognitive accuracy can also be assessed during other types of memory tasks (such as those that are recognition rather than recall based) and during tasks that do not have a strong memory component.

In the current study, we operationalized relative metacognitive accuracy as the gamma correlation coefficient. While doing so allows us to better compare the results of this study to previous work conducted in the metacognition literature, correlational measures of metacognitive accuracy such as gamma and phi coefficients are not without controversy. For example, gamma correlations can be constrained by metacognitive bias, such that gamma correlations for individuals with a liberal response criterion (and higher false alarm rates) will show lower relative accuracy regardless of actual metacognitive discrimination ability (Masson & Rotello, 2009). Age differences in these biases can make it difficult to draw conclusions about age-related changes in metacognitive accuracy using gamma. Although there are measures of metacognitive accuracy that are more robust to the effects of bias (e.g. *d*' and meta-*d*'; Fleming & Lau, 2014), these measures are traditionally computed for recognition memory tasks rather than recall tasks such as the one used in the current study.

Finally, an additional potential limitation was the control condition that was used in the current study. A goal of the current study was to examine reactivity associated with having participants make metacognitive judgements. To accomplish this we compared monitoring frequency, monitoring accuracy, and task performance in a group that made JOLs and a group that was not required to provide any type of judgement regarding the task stimuli. While participants in were equated in how long they viewed and could study the paired associates, it is possible that we would have observed a different pattern of results had we used a different control condition. One avenue for future research would be to include a control condition where participants are still required to provide some type of judgement during the paired associates task. For example, we might ask control participants to provide judgements regarding the emotional valence of the study items. Including a control condition that is still required to provide some type of judgment regarding the task stimuli could allow us to determine with greater certainty whether increases in monitoring frequency and recall performance within our experimental condition were the result of having them provide JOLs. It is possible that participants would report more monitoring and obtain higher recall performance even if they provided non-metacognitive judgements during the task.

Older adults often demonstrate spared monitoring ability. In the current study, older adults showed comparable JOL resolution to younger adults. Additionally, older adults in the current study had paired associates task recall scores that were comparable to younger adults and older adults also reported using as many effective strategies to learn the word pairs as younger adults did. The results of the study support previous findings that older adults demonstrate spared monitoring ability, although increased monitoring frequency does not seem to account for this age-invariance in monitoring ability. If age-related increases in propensity to engage in monitoring do not explain age-invariance in monitoring accuracy, then additional follow up studies are needed to determine why and how older adults often monitor as accurately as do younger adults.

REFERENCES

Andrews-Hanna, J.R., Reidler, J.S., Huang, C., & Buckner, R.L. (2010). Evidence for the default network's role in spontaneous cognition. *Journal of Neurophysiology, 104*(1), 322–335. doi: 10.1152/jn/00830.2009

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

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging. The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes. (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). New York: Cambridge University Press.

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. doi:10.1002/acp.1553

Beadel, J.R., Green, J.S., Hosseinbor, S., & Teachman, B.A. (2013). Influence of age, thought, content, and anxiety on suppression of intrusive thoughts. *Journal of Anxiety Disorder, 27*(6), 598-607. doi: 10.1016/j.janxdis.2012.12.002

Beaman, C. P., & Williams, T. I. (2010). Earworms (stuck song syndrome): Towards a natural history of intrusive thoughts. *British Journal of Psychology*, *101*(4), 637 653.doi: 10.1348/000712609X4 79636

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. doi:10.1016/0749596X(89)90016-8

Berntsen, D., & Rubin, D.C. (2002). Emotionally charged autobiographical memories across the lifespan: The recall of happy, sad, traumatic, and involuntary memories. *Psychology and Aging, 17*(4), 636-652. doi: 10.1037/0882 7974.17.4.636

Brigham, M.C. & Pressley, M. (1988). Cognitive monitoring and strategy choice in younger and older adults. *Psychology and Aging, 3*(3), *3*49–257. doi: 10.1037/0882-7974.3.3.249

Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*, 1-16. doi: 10.1111/j.2044-8260.1982.tb01421.x

Bruce, P. R., Coyne, A. C., & Botwinick, J. (1982). Adult age differences i metamemory. *Journal of Gerontology, 37*(3), *3*54–357. doi: 10.1093/geronj/37.3.354

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

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. doi: 10.3758/BF03194036

Christoff, K., Gordon, A.M., & Smith, R. (2011). The role of spontaneous thought in human cognition. In O. Vartanian & D.R. Mandel (Eds.), *Neuroscience of decision making* (pp. 259–284). New York, NY: Psychology Press.

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. doi: 10.1037/08827974.12.1.50

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. doi: 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. doi: 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.

Crovitz, H. F., & Daniel, W. F. (1984). Measurements of everyday memory: Toward the prevention of forgetting. *Bulletin of Psychonomic Society*, *22*(5), 413–414. doi: 10.3758/BF03333861

Devolder, P. A., Brigham, M. C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging, 5*(2), 291–303. doi:10.1037/08827974.5.2.291

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. doi: 10.1037/08827974.22.1.122

Double, K.S., Birney, D.P., & Walker, S.A. (2017). A meta-analysis and systematic review of reactivity to judgements of learning. *Memory*, *21*(1-10). doi: 0.1080/09658211.2017.1404111.

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. doi: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. doi:10.1037//0882 7974.15J.4

Dunlosky, J., & Hertzog, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multi-trial learning. *Journal of*

Gerontology: Psychological Sciences, 52(4), 178-186. doi: 0.1093/geronb/52B.4.P178

Dunlosky, J., Hertzog, C., Kennedy, M., & Thiede, K. (2005). The self-monitoring approach for effective learning. *Cognitive Technology, 10*(1), 4-11.

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. doi: 10.1080/13825585.2013.775217

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods*, *39*,175-191.

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

Fleming, S. M. & Lau, H. C. (2014) How to measure metacognition. *Frontiers in Human Neuroscience, 8*(443). doi: 10.3389/fnhum.2014.00443.

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. doi: 10.1037/pag0000031

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

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*(268), 732–764.

Halamish, V., McGillivray, S., & Castel, A. D. (2011). Monitoring one's own forgetting in younger and older adults. *Psychology and Aging, 26*(3), 631–635. doi:10.1037/a0022852

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

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (vol. 22, pp. 193-225). San Diego, CA: Academic Press.

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. doi: 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. doi: 10.1037/0882-7974.5.2.215

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. doi: 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), *6*07-624. doi: 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. doi: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. doi: 10.1037/0882-7974.17.2.209

Jackson, J. D., and Balota, D. A. (2012).Mind wandering in younger and older adults: converging evidence from the sustained attention to response task and reading for comprehension. *Psychology & Aging, 27,* 106–119.

Jordano, M. J., & Touron, D. R. (2017a). Stereotype threat as a trigger of mind wandering in older adults. *Psychology and Aging.* doi: 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. doi: 10/1016/j.concog.2017.08.002.

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. doi: 10.1016/S0749596X(02)00504-1

Kopp, K., D'Mello, S., & Mills, C. (2015). Influencing the occurrence of mind wandering while reading. *Consciousness & Cognition*, *34*, 52-62.

Koriat, A. (1981). [Confidence and self-consistency: Evidence from within-person analyses]. Unpublished raw data.

Krawietz, S. A., Tamplin, A. K., & Radvansky, G. A. (2012). Aging and mind wandering during text comprehension. *Psychology and Aging, 27*(4), 951–958. doi: 10.1037/a0028831

Kuhlmann, B. G., & Touron, D. R. (2011). Older adults' use of metacognitive knowledge in source monitoring: Spared monitoring but impaired control. *Psychology and Aging, 26*(1),143–149.doi:10.1037/a0021055

Kvavilashvili, L., & Fisher, L. (2007). Is time-based prospective remembering mediated by self initiated rehearsals?: Role of cues, ongoing activity, age and motivation. *Journal of Experimental Psychology: General, 136*(1), 112-132. doi: 10.1037/0096-3445.136.1.112

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology, 48*(1), 47-94. doi: 10.1016/S0010 0285(03)00115-4

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. doi: 10.3758/s13421-011-0140-x

Lambert, A.E., Smyth, F.L., Beadel, J.R., & Teachman, B.A. (2013). Aging and repeated thought suppression success. *PLoS One*, *8*(6), e65009. doi: 10.1371/ journal.pone.0065009

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. doi: 10.1037/0278 7393.16.3.464

Levy-Cushman, J., & Abeles, N. (1998). Memory complaints in the able elderly. *Clinical Gerontologist, 19*(2), 3-24. doi: 10.1300/J018v19n02_02

Loewen, E.R., Shaw, R.J. & Craik, F.I.M. (1990). Age differences in components of metamemory. *Experimental Aging Research, 16*(1), 43-48. doi: 10.1080/03610739008253873

Lovelace, E. A. 1990. "Aging and metacognitions concerning memory function". In E.A. Lovelace (Ed.) *Aging and cognition: Mental processes, self awareness, and interventions* (pp. 157–188). Amsterdam: North Holland.

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. doi: 10.1093/geronj/40.2.192

Magee, J.C., & Teachman, B.A., (2012). Distress and recurrence of intrusive thoughts in younger and older adults. *Psychology and Aging, 27*(1), 199-210. doi: 10.1037/a0024249

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

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology, 97*(4), 723-731.

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35,* 509–527. doi: 10.1037/a0014876

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.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14,* 127–144. doi: 10.1002/acp.775

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. doi: 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. doi: 10.1037/a0014104.

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "D'oh!": working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(3), 525-549. doi:0.1037/a0025896

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. doi: 10.1016/j.actpsy.2012.11.006

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

Moulin, C.J.A., Souchay, C., Buchanan, S., Bradley, R., Karadoller, D.Z., & Akan, M. (2014). Déjà Vu in Older Adults. In: B.L. Schwartz & A.S. Brown (Eds.). *Tip-of the-tongue states and related phenomena* (pp. 281-304). New York, NY: Cambridge University Press.

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. doi: 10.1016/j.jesp.2011.05.011

Mullet, H.G., Scullin, M.K., Hess, T.J., Scullin, R.B., Arnold, K.M., & Einstein, G.O. (2013). Prospective memory and aging: evidence for preserved spontaneous retrieval with exact but not related cues. *Psychology and Aging*, *28*(4), 910-22. doi: 10.1037/a0034347

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of knowing predictions. *Psychological Bulletin*, *95*(1), 109-133. doi: 10.1037/0033 2909.95.1.109

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267–270. doi:10.1111/j.1467-9280.1991.tb00147.x

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. doi: 10.1037/0278-7393.14.4.676

Nelson, T. O., & Narens, L. (1990). Metamemory: a theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). New York: Academic Press.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

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. doi: 10.1037/0278-7393.8.4.279

O'Callaghan, C., Shine, J.M., Lewis, S.J., Andrews-Hanna, J.R., & Irish, M. (2015). Shaped by our thoughts – A new task to assess spontaneous cognition and its associated neural correlates in the default network. *Brain and Cognition, 93,* 1 10. doi:10.1016/j.bandc.2014.11.001

Pansky, A., Goldsmith, M., Koriat, A., & Pearlman-Avnion, S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology, 21*(2-3), 303–329. doi:10.1080/09541440802281183

Pressley, M., Levin, J.R., & Ghatala, E.S. (1984). Memory strategy monitoring in adults and children. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 270-288. doi:10.1016/S0022-5371(84)90181-6

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. doi: 10.1093/geronj/37.6.688

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

Rubin, D.C., & Berntsen, D. (2009). The frequency of voluntary and involuntary autobiographical memories across the life span. *Memory and Cognition*, *37*(5), 679-688. doi: 10.3758/37.5.679

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.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition & Learning, 4,* 33-45

Schmader, T. (2010). Stereotype threat deconstructed. *Current Directions in Psychological Science, 19*(1), 14-18.

Scullin, M.K., Bugg, J.M., McDaniel, M.A., & Einstein, G.O. (2011). Prospective memory and aging: preserved spontaneous retrieval, but impaired deactivation, in older adults. *Memory and Cognition*, *39*(7), 1232-40. doi: 103758/s13421-011 0106-z

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*(8), 605-617. doi:10.1016/j.tics.2016.05.010

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. doi: 0882-7974/89/800.75

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

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annual Review of Psychology, 66,* 487 518. doi: 10.1146/annurev-psych-010814-015331

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. doi: 10.1037/a0014855

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. doi: 10.1016/j.concog.2004.06.003

Souchay, C., & Isingrini, M. (2012). Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. *Acta Psychologica, 139*(3), 458-64. doi: 10.1016/j.actpsy.2012.01.007

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

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, & Cognition*, *39*, 832-842.

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. doi: 10.1016/j.intell.2009.08.002

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.

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. doi:10.1016/j.actpsy.2013.10.016

FOOTNOTES

1. Because we would not consider off-task thoughts about length of task to be an example of metacognitive monitoring, we have excluded "other" TRI from our TRI variable used in these analyses. Thus, the TRI variable included in these analyses consisted of approach-based and evaluative TRI. Excluding "other" TRI from our TRI variable did not influence the patterns of results obtained.

APPENDIX A

DSSQ (THOUGHT CONTENT)

This set of questions concerns the kinds of thoughts that go through people's heads at particular times, for example while they are doing some task or activity. Below is a list of thoughts, some of which you might have had recently. Please indicate roughly how often you had each thought while performing the task, by circling a number from the list below.

1= Never 2= Once 3= A few times 4= Often 5= Very often

1. I thought about how I should work more carefully. 1 2 3 4 5
2. I thought about how much time I had left. 1 2 3 4 5
3. I thought about how others have done on this task. 1 2 3 4 5
4. I thought about the difficulty of the problems. 1 2 3 4 5
5. I thought about my level of ability. 1 2 3 4 5
6. I thought about the purpose of the experiment. 1 2 3 4 5
7. I thought about how I would feel if I were told how I performed. 1 2 3 4 5
8. I thought about how often I get confused. 1 2 3 4 5
9. I thought about members of my family. 1 2 3 4 5
10. I thought about something that made me feel guilty. 1 2 3 4 5
11. I thought about personal worries. 1 2 3 4 5
12. I thought about something that made me feel angry. 1 2 3 4 5
13. I thought about something that happened earlier today.1 2 3 4 5
14. I thought about something that happened in the recent past 1 2 3 4 5
(last few days, but not today).
15. I thought about something that happened in the distant past 1 2 3 4 5
16. I thought about something that might happen in the future. 1 2 3 4 5

Scoring: Items 1-18 on this questionnaire are summated to provide a measure of task-related interference (TRI). Items 9-16 on this questionnaire are summated to provide a measure of task-unrelated thinking (TUT).

APPENDIX B

DSSQ (MOTIVATION)

Dundee Stress State Questionnaire (Motivation)

Please answer some questions about your attitude to the task you are about to do. Rate your agreement with the following statements by circling one of the following answers:

Extremely = 4 Very much = 3 Somewhat = 2 A little bit = 1 Not at all = 0

1. The content of the task was interesting 0 1 2 3 4
2. The only reason to do the task is to get an external reward (e.g. payment) 0 1 2 3 4
3. I would rather have spent the time doing the task on something else 0 1 2 3 4
4. I was concerned about not doing as well as I can 0 1 2 3 4
5. I wanted to perform better than most people do 0 1 2 3 4
6. I became fed up with the task 0 1 2 3 4
7. I was eager to do well 0 1 2 3 4
8. I would be disappointed if I failed to do well on this task 0 1 2 3 4
9. I was committed to attaining my performance goals0 1 2 3 4
10. Doing the task was worthwhile 0 1 2 3 4
11. I found the task boring 0 1 2 3 4
12. I felt apathetic about my performance 0 1 2 3 4
13. I wanted to succeed on the task 0 1 2 3 4
14. The task brought out my competitive drives 0 1 2 3 4
15. I was motivated to do the task 0 1 2 3 4

Scoring: Items 1-15 assess motivation. Items 4, 5, 7, 8, 9, 13, and 14 are summated to get a measure of success motivation, or motivation to excel in task performance. Item number 15 can be used to provide an overall level of motivation, if needed. The remainder of the items are summated to provide a measure of how interesting participants thought the task was. Items 1 and 10 are positively scores while items 2, 3, 6, 11, and 12 are reversed scored.

APPENDIX C

STUDIES 1-5 POST-TASK QUESTIONS

1. If you reported thinking about task approach on the mind-wandering probes, what were you thinking about? Please provide an explanation or example below:

2. If you reported thinking about task evaluation on the mind-wandering probes, what were you thinking about? Please provide an explanation or example below:

3. If you reported thinking about task approach on the mind-wandering probes, to what extent were those thoughts worry-laden?

   1=Not at all worry-laden and 5=Very worry-laden

4. If you reported thinking about task evaluation on the mind-wandering probes, to what extent were those thoughts worry-laden?

   1=Not at all worry-laden and 5=Very worry-laden

5. If you used strategies to remember letters during the task, please provide an example of a strategy you used:

6. If you used strategies to verify math problems during the task, please provide an example of a strategy you used:

7. Please rate the effectiveness of the strategy you described that helped you to remember letters during the task.

   1=Not at all effective and 5=Very effective

8. Please rate the effectiveness of the strategy you described that helped you to verify math problems during the task.

   1=Not at all effective and 5=Very effective

9. How difficult did you find the experimental task overall?

   1=Not at all difficult and 5= Very difficult

10. How difficult did you find the math portion of the experimental task?

    1=Not at all difficult and 5= Very difficult

11. How difficult did you find the letter recall portion of the experimental task?

    1=Not at all difficult and 5= Very difficult

12. How well do you think you did in this study overall?

    1=Very poorly and 5= Very well

13. Does your performance on the experimental task satisfy you?

    1=Not at all and 5= Very much

14. How much effort did you put into this study?

    1=Very little effort and 5= A lot of effort

15. Would you be able to do better in this study if you tried harder?

    1=Very unlikely and 5= Very likely

16. Were you interested in this study?

    1=Not at all interested and 5=Very interested

17. How fatiguing did you find the experimental task overall?

    1=Not at all fatiguing and 5=Very fatiguing

18. Did you feel stress or tension during this study?

    1=Not at all and 5= Very much

19. Did you have to read instructions on the screen multiple times before you understood them?

    1=Never and 5= Always

20. Did you get distracted easily during the study?

    1=Not at all and 5= Very much

21. How focused were you on accurately recalling letters in the correct order during the experimental task?

    1=Not at focused and 5= Very focused

22. How focused were you on accurately verifying the math equations during the experimental task?

    1=Not at focused and 5= Very focused

APPENDIX D

PTQ DESCRIPTIVE STATISTICS FOR STUDY 1

Measure	M	SE
PTQ Experiment Difficulty	3.23	.12
PTQ Math Difficulty	3.47	.11
PTQ Letter Recall Difficulty	3.10	.15
PTQ Well	3.47	.10
PTQ Satisfy	3.17	.11
PTQ Motivation	3.60	.12
PTQ Better	2.70	.13
PTQ Interest	3.43	.10
PTQ Fatigue	3.40	.12
PTQ Stress	3.52	.14
PTQ Understand	2.60	.14
PTQ Distract	2.60	.14
PTQ Letter Focus	3.87	.10
PTQ Math Focus	3.57	.11
PTQ Stereotype	3.40	.12
DSSQ TUT	14.83	.92
DSSQ TRI	18.97	.92
PTQ Letter Strategy Effectiveness	4.13	.08
PTQ Math Strategy Effectiveness	3.87	.13
PTQ Task evaluation worry	2.62	.10
PTQ Task approach worry	2.42	.11

Note. All responses to the post-task questions above were made on a 1 to 5 Likert scale. For PTQ Overall Task Difficulty, PTQ Math Difficulty, and PTQ Recall Difficulty 1 = Not at all difficult and 5 = Very difficult. For PTQ Task Performance 1 = Very poorly and 5 = Very well. For all remaining questions 1 = Not at all and 5 = Very much. For PTQ Letter Strategy Effectiveness and PTQ Math Strategy Effectiveness 1 = Not at all effective and 5 = Very effective. For PTQ Task evaluation worry and PTQ task approach worry 1 = Not at all worry-laden and 5 = Very worry-laden.

APPENDIX E

PTQ DESCRIPTIVE STATISTICS FOR STUDY 2

	Non-branching		Branching		
Measure	*M*	*SE*	*M*	*SE*	*p*
PTQ Experiment Difficulty	3.23	.12	3.13	.14	.590
PTQ Math Difficulty	3.47	.11	2.87	.18	.006
PTQ Letter Recall Difficulty	3.10	.15	3.37	.17	.238
PTQ Well	3.47	.10	3.97	.11	.001
PTQ Satisfy	3.17	.11	2.73	.11	.006
PTQ Motivation	3.60	.12	4.00	.16	.050
PTQ Better	2.70	.13	2.80	.13	.589
PTQ Interest	3.43	.10	3.27	.15	.378
PTQ Fatigue	3.40	.12	3.67	.11	.103
PTQ Stress	3.52	.14	3.80	.17	.209
PTQ Understand	2.60	.14	3.10	.13	.011
PTQ Distract	2.60	.13	3.10	.12	.007
PTQ Letter Focus	3.87	.10	3.63	.12	.130
PTQ Math Focus	3.57	.11	3.73	.16	.413
PTQ Stereotype	3.40	.12	3.43	.12	.860
DSSQ TUT	14.83	.92	15.97	1.01	.407
DSSQ TRI	18.97	.92	20.30	.86	.295
PTQ Letter Strategy Effectiveness	4.13	.08	4.17	.10	.756
PTQ Math Strategy Effectiveness	3.87	.13	3.60	.17	.212
PTQ Task evaluation worry	2.57	.10	2.60	.12	.848
PTQ Task approach worry	2.62	.11	2.53	.12	.582

Note. All responses to the post-task questions above were made on a 1 to 5 Likert scale. For PTQ Overall Task Difficulty, PTQ Math Difficulty, and PTQ Recall Difficulty 1 = Not at all difficult and 5 = Very difficult. For PTQ Task Performance 1 = Very poorly and 5 = Very well. For all remaining questions 1 = Not at all and 5 = Very much. For PTQ Letter Strategy Effectiveness and PTQ Math Strategy Effectiveness 1 = Not at all effective and 5 = Very effective. For PTQ Task evaluation worry and PTQ task approach worry 1 = Not at all worry-laden and 5 = Very worry-laden.

APPENDIX F

PTQ DESCRIPTIVE STATISTICS FOR STUDY 3

	Non-branching		Branching		
Measure	*M*	*SE*	*M*	*SE*	*p*
PTQ Experiment Difficulty	3.23	.12	3.13	.14	.590
PTQ Math Difficulty	3.47	.11	2.87	.18	.006
PTQ Letter Recall Difficulty	3.10	.15	3.37	.17	.238
PTQ Well	3.47	.10	3.97	.11	.001
PTQ Satisfy	3.17	.11	2.73	.11	.006
PTQ Motivation	3.60	.12	4.00	.16	.050
PTQ Better	2.70	.13	2.80	.13	.589
PTQ Interest	3.43	.10	3.27	.15	.378
PTQ Fatigue	3.40	.12	3.67	.11	.103
PTQ Stress	3.52	.14	3.80	.17	.209
PTQ Understand	2.60	.14	3.10	.13	.011
PTQ Distract	2.60	.13	3.10	.12	.007
PTQ Letter Focus	3.87	.10	3.63	.12	.130
PTQ Math Focus	3.57	.11	3.73	.16	.413
PTQ Stereotype	3.40	.12	3.43	.12	.860
DSSQ TUT	14.83	.92	15.97	1.01	.407
DSSQ TRI	18.97	.92	20.30	.86	.295
PTQ Letter Strategy Effectiveness	4.13	.08	4.17	.10	.756
PTQ Math Strategy Effectiveness	3.87	.13	3.60	.17	.212
PTQ Task evaluation worry	2.57	.10	2.60	.12	.848
PTQ Task approach worry	2.62	.11	2.53	.12	.582

Note. All responses to the post-task questions above were made on a 1 to 5 Likert scale. For PTQ Overall Task Difficulty, PTQ Math Difficulty, and PTQ Recall Difficulty 1 = Not at all difficult and 5 = Very difficult. For PTQ Task Performance 1 = Very poorly and 5 = Very well. For all remaining questions 1 = Not at all and 5 = Very much. For PTQ Letter Strategy Effectiveness and PTQ Math Strategy Effectiveness 1 = Not at all effective and 5 = Very effective. For PTQ Task evaluation worry and PTQ task approach worry 1 = Not at all worry-laden and 5 = Very worry-laden.

APPENDIX G

DEMOGRAPHICS AND HEALTH QUESTIONNAIRE

Participant Information Survey

In order to better understand the results of the study you have agreed to participate in, we need to know a few things about you and your background.
We will use this information for research purposes only, and it will be kept strictly confidential. You will note that we do not ask for your name during this survey.
Please respond to the following questions completely. Ask the Experimenter if you need assistance in answering any question.
If you have limited experience with computers, or are unsure about how to use the computer in answering these questions, please ask the Experimenter for assistance at any time.
Please press ENTER to begin.

1.My sex is (please circle):MaleFemale

2.My birth date is:

(Month)   (Day)(Year)

3. What is your native language?

English: _____Other (please specify): ________________

4.What is your ethnic background? Please check the appropriate alternative.

_____**First Nations** origin (A person having origins in any of the original peoples of North America, and who maintains a cultural identification through tribal or band affiliation or community recognition)

_____**Asian or Pacific Islander** (A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Pakistan, Japan, Korea, the Philippine Islands, and Samoa.)

_____**Black**, not of Hispanic origin (A person having origins in any of the black racial groups of Africa)

_____**Hispanic** (A person of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish culture or origin, regardless of race)

______**White**, not of Hispanic origin (A person having origins in any of the original peoples of Europe, North Africa, or the Middle East)

5. Which academic diplomas / degrees / certificates have you obtained? (Please circle **ALL** that apply)

a) no degree
b) high school diploma
c) technical/trade school or community college
d) Bachelor’s (e.g., BA, BSc, BComm.)
e) Master’s (e.g., MA, MSc, MEd, LLM)
f) Law degree (Bachelor=s, LLB)
g) Medical degree (MD)
h) PhD or other doctoral degree
i) other or additional degrees (please specify)

6. For **EACH** of the following levels of education, please circle the highest grade or years of full-time attendance you have **COMPLETED.** Do not include part-time or extension courses taken for interest.

a) <u>Grade/Intermediate School</u>

Grade 1   Grade 2   Grade 3   Grade 4   Grade 5   Grade 6   Grade 7   Grade 8

b) <u>Secondary/High School</u>

none Grade 9   Grade 10   Grade 11   Grade 12   Grade 13

c) <u>Technical, Trade, Nursing or Business School, or Community College</u>

none 1 year 2 years 3 years 4 years 5+ years

d) University (Bachelor=s Level)

none 1st year 2nd year 3rd year 4th year 5th year

e) Post-Graduate School (e.g., Master's, PhD)

none 1 year 2 years 3 years 4 years 5+ years

7.Are you currently involved in volunteer work? Yes _____No _____

If yes, please briefly describe your volunteer activities:

____

8. Are you currently a student? Yes _____ No _____

If yes, how many hours a week do you spend in classes? _______ hrs

Are you pursuing a specific certificate, diploma or degree? Yes _____ No _____

Please briefly describe what you are studying:

____

9. Compared to a perfect state of health, I believe my overall health to be (Please circle one):

a. very good
b. good
c. fair
d. poor
e. very poor

10. Compared to other people my age, I believe my overall health to be (Please circle one):

a. very good
b. good
c. fair
d. poor
e. very poor

11. Compared to other people my age, I believe my eyesight to be (Please circle one):

a. very good
b. good
c. fair
d. poor
e. very poor

12. Compared to other people my age, I believe my hearing to be (Please circle one):

a. very good
b. good

c. fair
d. poor
e. very poor

13. In the past 3 years, my health has affected my daily activities in the following way (Please circle one):
a. not applicable
b. improved
c. no change
d. slightly reduced

e. moderately reduced

f. drastically reduced

g. gave up employment

h. gave up travel

14. The following chart lists a number of health-related conditions that may apply to you. Please answer all parts of each question as concisely as you can. The first part of each question is the most important for you to answer. Specifically, we would like to know whether or not you have ever been <u>diagnosed by a medical practitioner</u> with the condition in question. **If you responded YES,** please complete the remaining parts of the question.

	If YES, how serious is your condition? (please check one)				**At what AGE were you diagnosed with this condition?**
<u>Do you suffer from this condition?</u>	**No**	**Yes, not serious**	**Yes, moderately serious**	**Yes, very serious**	
Hearing problems (e.g., tinitus) that cannot be corrected with a hearing aid					
Visual disorders (e.g., glaucoma, cataracts, macular degeneration) that cannot be corrected with glasses					
Asthma					

Bronchitis					
Tuberculosis					
Hardening of the arteries					
High blood pressure					
Stroke					
Low blood pressure					
Gall bladder problems					
Liver trouble					
Stomach ulcer					
Kidney or bladder trouble or cystitis					
Gynecological problems					
Colitis or diverticulitis					
Paralysis not related to stroke					
Spinal condition (e.g., scoliosis)					
Back trouble					
Parkinson's disease					
Epilepsy					
Thyroid					
Prostate problems					
Anemia					
Depression					
Alcohol dependence					
Drug dependence					
Heart trouble					
Osteo-arthritis					
Rheumatoid arthritis					
Osteoporosis					

Diabetes (sugar sickness)					
Cancer					
Migraine					
Encephalitis					
Meningitis					
Head injury					

15. Are you presently taking any drugs or medications (prescription or other)?

a. Yes
b. No

If yes, here is a chart of medications that people often have to take. Please indicate whether you are taking any of these medications (Check all that apply).

	medicine for high blood pressure (e.g., Prinivil, Lopressor, Procardia, Vasotec, etc.)
	digitalis or other medication for your heart
	medicine for chest pain/angina (e.g., Nitroglycerine, Digoxin, Procardia, etc.)
	any sort of diabetes medicine (pills, pumps, or injections: e.g., Glucotrol, Tolinase, Insulin)
	cortisone or anti-inflammatory drugs for arthritis (e.g., Prednisone, Tolectin, etc.)
	pills to make you lose water or salt (diuretics: e.g., Lasix, Bumex, etc.)
	tranquillizers or sedatives (e.g., Ativan, Xanax, Valium, etc.)
	sleeping pills/hypnotics (e.g., Chloral hydrate, Restoril, Dalmane, etc.)
	blood thinner medicine (anticoagulants: e.g., Coumadin, Heparin, etc.)
	vitamin or mineral supplements (e.g., Iron, Calcium, Potassium, etc.)
	female hormone supplements (e.g., Estrogen, Premarin, etc.)
	appetite suppressants or diet pills
	pain medication (more than 2-3 times a week: e.g., Tylenol, Advil, Percocet, Darvocet-N-100, etc.)
	allergy or asthma medicine
	ulcer or other stomach medicine (e.g., Tagamet, Lactate, Prilosec, etc)
	antibiotics (e.g., Penicillin, Ampicillin, Tetracycline, etc.)
	medicine to control seizures (e.g., Dilantin, Tegretol, etc.)
	medicine to control tremors (e.g., L-dopa, Sinemet, Parlodel, etc.)
	oral contraceptives
	stimulants to help you stay awake
	eye medication (e.g., eye drops/ointments: e.g., IsoptoCarpine, etc.)
	anti-depressant medication (e.g., Wellbutrin, Elavil, Zoloft, Prozac, etc.)
	anti-psychotic medication (e.g., Lithium, Prolixin, etc.)

	chemotherapy for cancer
	oral medication for cancer (e.g., Nolvadex, Cytoxan, etc.)
	medicine for a thyroid condition (e.g., Synthroid, Eltroxin, etc.)
	other prescription or non-prescription drugs (please indicate)

16. Today, have you taken any drugs or medications (prescription or other) that tend to make you drowsy?

a. Yes
b. No

17. Do you smoke or use tobacco products? (Please circle one)

a. Yes At what age did you start smoking? __________
b. No, I previously used tobacco but I have quit completely
For how many years did you use tobacco? __________
c. No, I have never used tobacco

If you **currently** use tobacco, what do you use? (Please circle. Complete all that apply)

a. **Cigarettes**
How many cigarettes do you smoke?
__________ cigarettes per __________ (day/week/month/year)
b. **Cigars**
How many cigars do you smoke?
__________ cigars per __________ (day/week/month/year)
c. **Pipe**
How many pipe bowls do you smoke?
__________ pipe bowls per __________ (day/week/month/year)
d. **Snuff or chewing tobacco**
How many pinches or plugs do you use?
__________ pinches/plugs per__________ (day/week/month/year)

18. Do you drink alcoholic beverages? (Please circle one)

a. Yes At what age did you start drinking? __________
b. No, I used to drink but have now completely given it up
For how many years did you drink? __________
c. No, I never drink

If you **currently** drink alcoholic beverages, what do you drink? (Please circle. Complete all that apply)

a. **Beer**
How many cans/bottles of beer do you consume?
__________ bottles/cans per __________ (day/week/month/year)
b. **Wine**
How many glasses of wine do you consume?

__________ glasses per __________ (day/week/month/year)

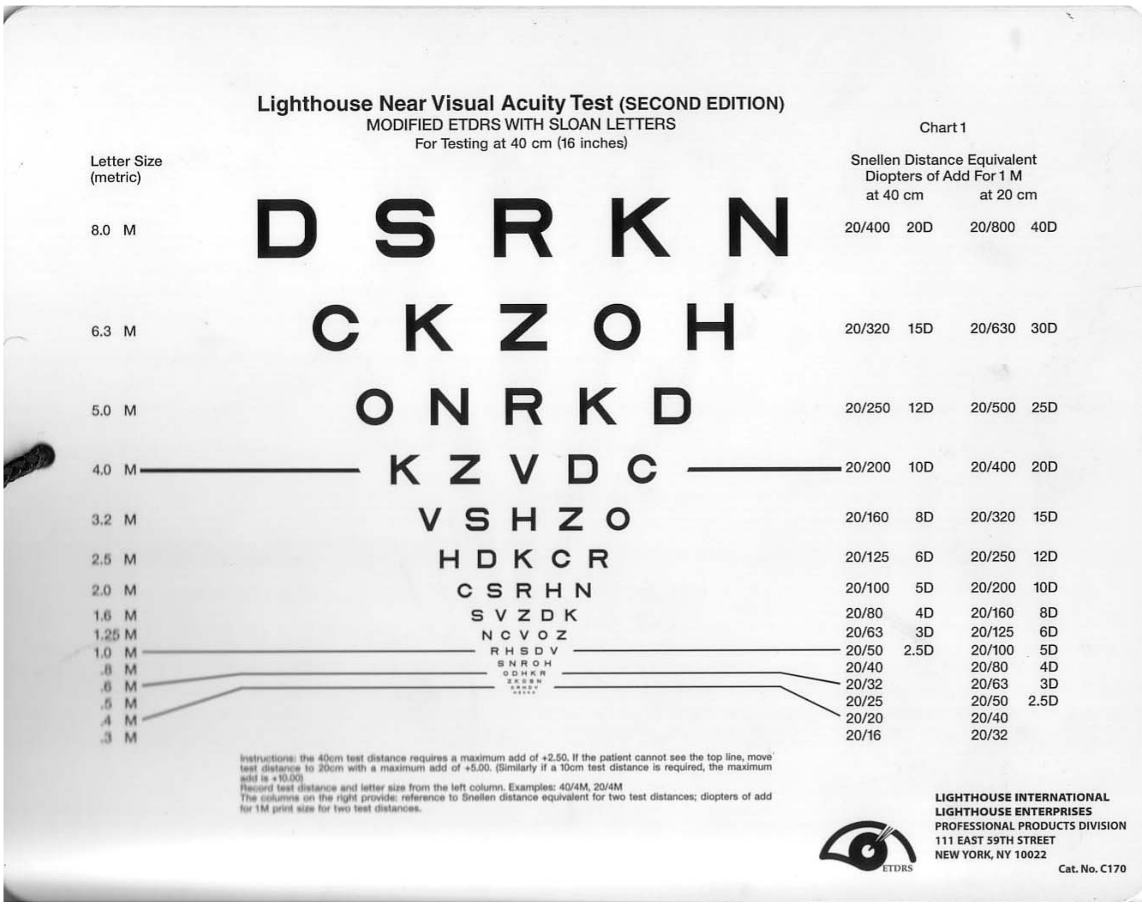
c. **Hard liquor** (i.e., with no mix added)
How many drinks do you consume? (1 drink = 1 ounce of alcohol)
__________ drinks per __________ (day/week/month/year)

d. **Mixed drinks** (i.e., alcohol with mix added)
How many mixed drinks do you consume? (1 drink = 1 ounce of alcohol)
__________ mixed drinks per __________ (day/week/month/year)

Thank You for your time

APPENDIX H

LIGHTHOUSE NEAR VISUAL ACUITY



Verbal instructions:

> "This is a near acuity test. Hold this card such that the end of the cord is next to your eye and the string is pulled taut. Now read the lowest line of which you can easily read all five letters."

Scoring:

> If a participant reads a line incorrectly they are instructed to read the line immediately above it. This process repeats until the participant in able to correctly read all five letters of a line. Scores are based on the lowest line that participants can read all five letter for.

APPENDIX I

ADVANCED VOCABULARY TEST I-V4

Name ____

ADVANCED VOCABULARY TEST I — V-4

This is a test of your knowledge of word meanings. Look at the sample below. One of the five numbered words has the same meaning or nearly the same meaning as the word above the numbered words. Mark your answer by putting an X through the number in front of the word that you select.

jovial

1-refreshing
2-scare
3-thickset
4-wise
X-jolly

The answer to the sample item is number 5; therefore, an X has been put through number 5.

Your score will be the number marked correctly minus a fraction of the number marked incorrectly. Therefore, it will not be to your advantage to guess unless you are able to eliminate one or more of the answer choices as wrong.

You will have 4 minutes for each of the two parts of this test. Each part has one page. When you have finished Part 1, STOP. Please do not go on to Part 2 until you are asked to do so.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

Part 1 (4 minutes)

1. mumble
   1-speak indistinctly
   2-complain
   3-handle awkwardly
   4-fall over something
   5-tear apart

2. perspire
   1-struggle
   2-sweat
   3-happen
   4-penetrate
   5-submit

3. gush
   1-giggle
   2-spout
   3-sprinkle
   4-hurry
   5-cry

4. massive
   1-strong and muscular
   2-thickly populated
   3-ugly and awkward
   4-huge and solid
   5-everlasting

5. feign
   1-pretend
   2-prefer
   3-wear
   4-be cautious
   5-surrender

6. unwary
   1-unusual
   2-deserted
   3-incautious
   4-sudden
   5-tireless

7. veer
   1-change direction
   2-hesitate
   3-catch sight of
   4-cover with a thin layer
   5-slide

8. orthodox
   1-conventional
   2-straight
   3-surgical
   4-right-angled
   5-religious

9. stripling
   1-stream
   2-narrow path
   3-engraving
   4-lad
   5-beginner

10. salubrious
    1-mirthful
    2-indecent
    3-salty
    4-mournful
    5-healthful

11. limpid
    1-lazy
    2-crippled
    3-clear
    4-hot
    5-slippery

12. procreate
    1-sketch
    2-inhabit
    3-imitate
    4-beget
    5-encourage

13. replete
    1-full
    2-elderly
    3-resentful
    4-discredited
    5-restful

14. frieze
    1-fringe of curls on the forehead
    2-statue
    3-ornamental band
    4-embroidery
    5-sherbet

15. treacle
    1-sewing machine
    2-framework
    3-leak
    4-apple butter
    5-molasses

16. ignominious
    1-inflammable
    2-elflike
    3-unintelligent
    4-disgraceful
    5-mysterious

17. abjure
    1-make certain
    2-arrest
    3-renounce
    4-abuse
    5-lose

18. duress
    1-period of time
    2-distaste
    3-courage
    4-hardness
    5-compulsion

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

STOP.

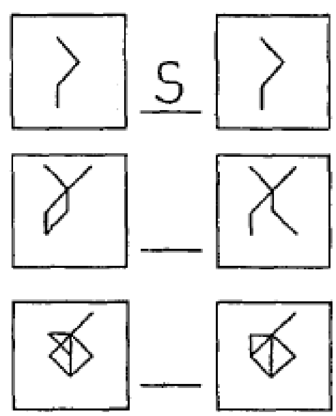
APPENDIX J

PROCESSING SPEED TASK

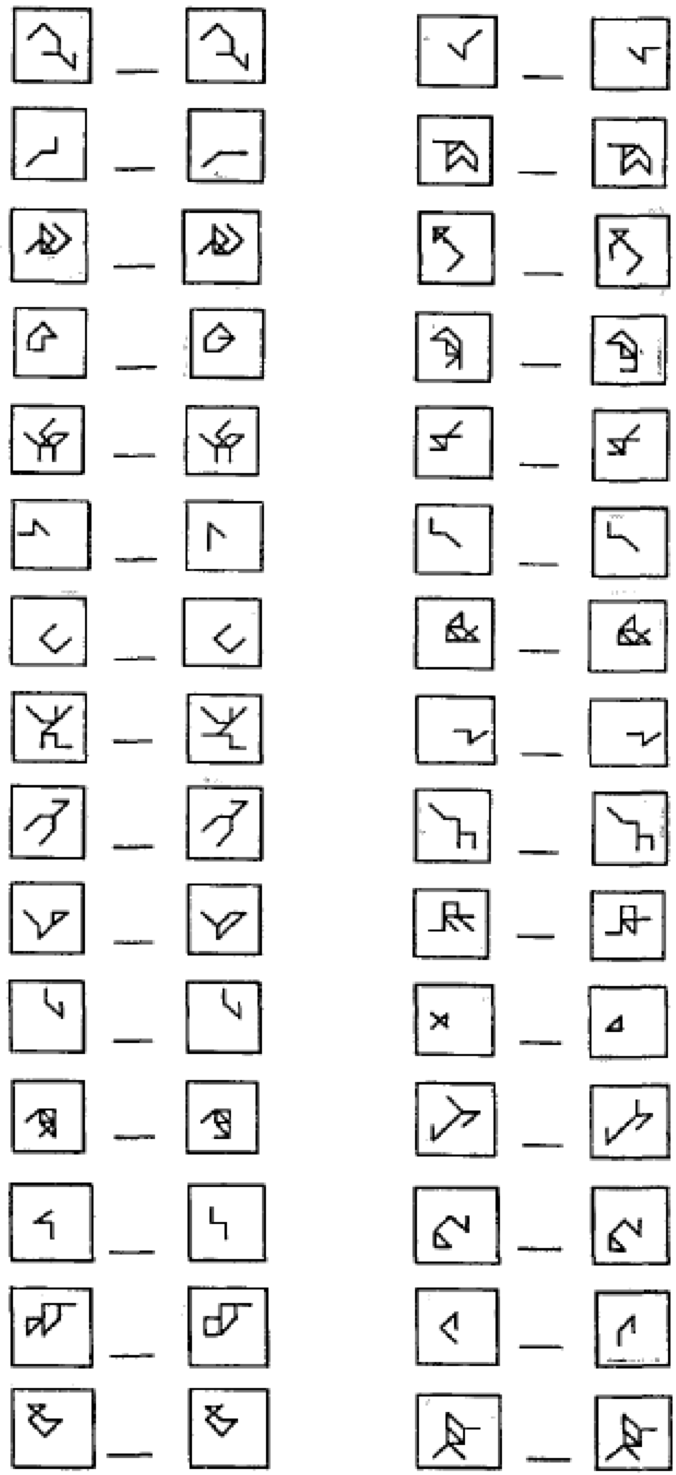
PATTERN COMPARISON

In this test you will be asked to determine whether two patterns of lines are the same or different. If the two patterns are the SAME, write an S on the line between them. If they are DIFFERENT, write a D on the line. Please try to work as rapidly as you can, writing an answer to each pair of line patterns.

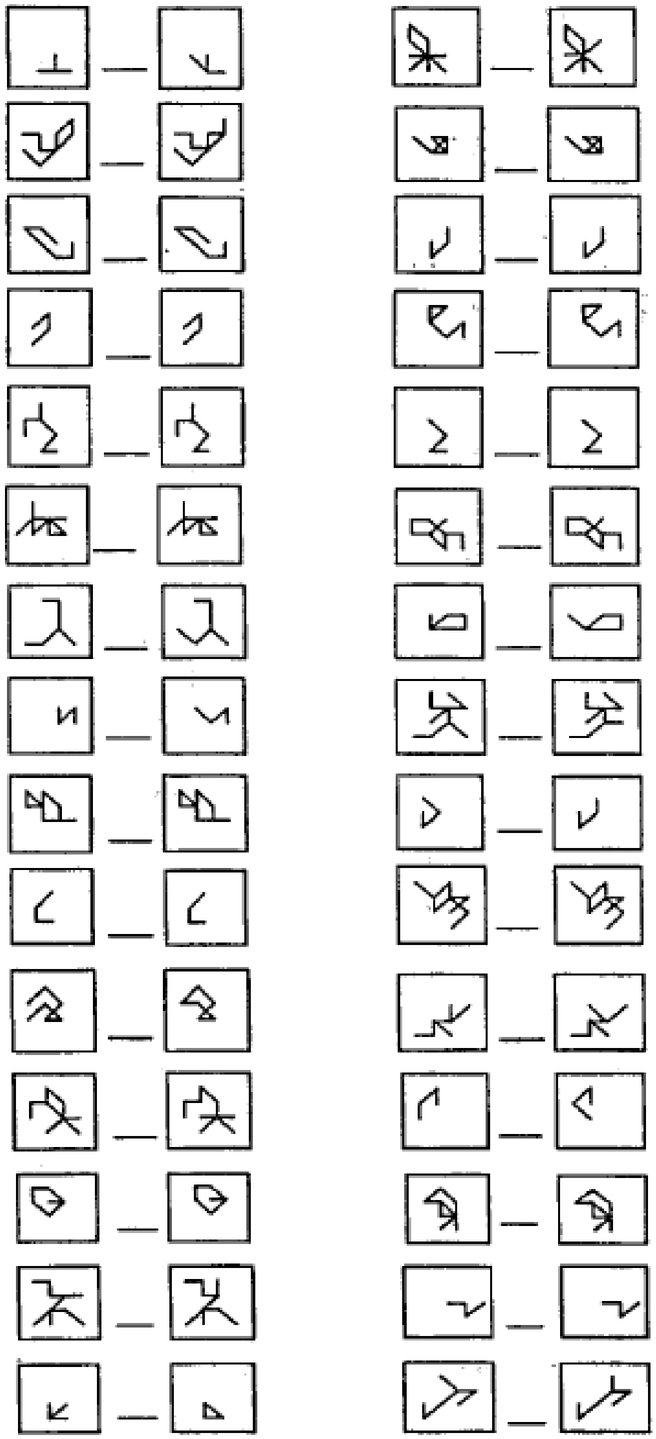
Try the following examples.



PLEASE DO NOT TURN THE PAGE UNTIL INSTRUCTED TO DO SO.



STOP



STOP

APPENDIX K

EXAMPLE STIMULI FOR PAIRED ASSOCIATES TASK

Unrelated pairs		Related pairs	
BONE	WINE	CHAIR	ARM
ANIMAL	LIBRARY	PAINT	ARTIST
THROAT	PRINCE	FLUTE	BAND
BABY	FOREST	PANTS	BELT
MESSAGE	SOLDIER	SKIN	BODY
MILK	CLERK	PAPER	BOOK
UNCLE	COAL	LEAF	BRANCH
COLUMN	KNIFE	CAB	BUS
TEXT	MAID	BICYCLE	CAR
TOOL	COAST	FENCE	CHAIN
ROAD	DANCER	HEART	CHEST
FLOWER	HORN	CAKE	PARTY
WHEEL	SENATOR	PIPE	CIGARETTE
WIRE	FORT	TIE	COAT
ROOM	BIRD	SUGAR	COFFEE
CROWD	RICE	PIG	COW
TONGUE	PICTURE	SIDEWALK	CONCRETE
HANDLE	CORN	LAMP	DESK
CAMERA	BRAIN	WALL	DOOR
TARGET	SEED	CRAYON	DRAWING
COTTON	SNAKE	SNOW	RAIN
DIRT	QUEEN	NOSE	EAR
COMPOSER	WOOD	MILL	FACTORY
DRESS	MOVIE	MIRROR	FACE
CLOTH	ATOM	OYSTER	FISH
PRISON	TISSUE	WEED	GARDEN
BAKER	WAGON	PLASTIC	GLASS
SALT	MAYOR	BEAR	HONEY
BEER	GRASS	SUIT	JACKET
TRAFFIC	BEACH	STOVE	KITCHEN

APPENDIX L

STUDY 5 POST-TASK QUESTIONS

1. How difficult did you find the memory task overall?
   1=Not at all difficult and 5= Very difficult

2. How well do you think you did in this study overall?
   1=Very poorly and 5= Very well

3. Does your performance on the memory test satisfy you?
   1=Not at all and 5= Very much

4. How much effort did you put into this study?
   1=Very little effort and 5= A lot of effort

5. How motivated were you to complete the memory test well?
   1=Very unmotivated and 5= Very motivated

6. Were you interested in this study?
   1=Not at all interested and 5=Very interested

7. How fatiguing did you find the experiment overall?
   1=Not at all fatiguing and 5=Very fatiguing

8. Did you feel stress or tension during this study?
   1=Not at all and 5= Very much

9. To what extent do you feel like you understood the task instructions?
   1=Not at all and 5= Very much

10. Did you get distracted easily during the study?
    1=Not at all and 5= Very much

APPENDIX M

COGNITIVE FAILURES QUESTIONNAIRE

We want to know how often these things have happened to you *in the past 6 months*. Please use the following scale:

0 Never	1 Very Rarely	2 Occasionally	3 Quite Often	4 Very Often

		Rating
1.	Do you read something and find you haven't been thinking about it and must read it again?	
2.	Do you find you forget why you went from one part of the house to the other?	
3.	Do you fail to notice signposts on the road?	
4.	Do you find you confuse right and left when giving directions?	
5.	Do you bump into people?	
6.	Do you find you forget whether you've turned off a light or a fire or locked the door?	
7.	Do you fail to listen to people's names when you are meeting them?	
8.	Do you say something and realize afterwards that it might be taken as insulting?	
9.	Do you fail to hear people speaking to you when you are doing something else?	
10.	Do you lose your temper and regret it?	
11.	Do you leave important letters unanswered for days?	
12.	Do you find you forget which way to turn on a road you know well but rarely use?	
13.	Do you fail to see what you want in a supermarket (although it's there)?	

14.	Do you find yourself suddenly wondering whether you've used a word correctly?	
15.	Do you have trouble making up your mind?	
16.	Do you find you forget appointments?	
17.	Do you forget where you put something like a newspaper or a book?	
18.	Do you find you accidentally throw away the thing you want and keep what you meant to throw away -- as in the example of throwing away the matchbox and putting the used match in your pocket?	
19.	Do you daydream when you ought to be listening to something?	
20.	Do you find you forget people's names?	
21.	Do you start doing one thing at home and get distracted into doing something else (unintentionally)?	
22.	Do you find you can't quite remember something although it's "on the tip of your tongue"?	
23.	Do you find you forget what you came to the shops to buy?	
24.	Do you drop things?	
25.	Do you find you can't think of anything to say?	

Scoring the Scale

The CFQ was developed by Broadbent et al. (1982) -- yes, the same Broadbent who proposed the filter theory of attention -- to assess the frequency with which people experienced cognitive failures, such as absent-mindedness, in everyday life -- slips and errors of perception, memory, and motor functioning. The most straightforward way to score the scale is simply to sum up the ratings of the 25 individual items, yielding a score from 0-100.

Scores on the scale predict episodes of absent-mindedness in both the laboratory and everyday life, including slow performance on focused attention tasks, traffic and work accidents, and forgetting to save one's data on the computer.

A study by Rast et al. (2008) indicates that the CFQ items load on three different factors. Summing scores across the relevant items will yield subscale scores representing these dimensions of forgetfulness:

- **Forgetfulness** (Items 1, 2, 5, 7, 17, 20, 22, and 23): "a tendency to let go from one's mind something known or planned, for example, names, intentions, appointments, and words".
- **Distractibility** (Items 8, 9, 10, 11, 14, 19, 21, and 25): "mainly in social situations or interactions with other people such as being absentminded or easily disturbed in one's focused attention".
- **False Triggering** (Items 2, 3, 5, 6, 12, 18, 23, and 24): "interrupted processing of sequences of cognitive and motor actions".

APPENDIX N

MEASURE OF DAILY MEMORY MONITORING

This set of questions concerns the kinds of thoughts that go through people's heads at particular times, for example while they are doing some task or activity. Below is a list of such thoughts. **Please indicate roughly how often you have each thought during when you are doing different things in your everyday life.**

1= Never 2= Once 3= A few times 4= Often 5= Very often

1. I think about how I should work more carefully.1 2 3 4 5
2. I think about how much time I had left.1 2 3 4 5
3. I think about how others have done on this task.1 2 3 4 5
4. I think about the difficulty of the problems.1 2 3 4 5
5. I think about my level of ability.1 2 3 4 5
6. I think about the purpose of the experiment.1 2 3 4 5
7. I think about how I would feel if I were told how I performed.1 2 3 4 5
8. I think about how often I get confused.1 2 3 4 5

This measure was adapted from the ‘Thinking Content’ scale of Sarason et al.'s (1986) Cognitive Interference Questionnaire (CIQ). The 8 items from this scale corresponding to task-related interference have been included in this measure to provide a rough measure of how much participants engage in thoughts related to monitoring in everyday life. Items 1-8 are summated to provide a measure of everyday monitoring or TRI.

APPENDIX O

STUDIES 1-4 TRI DATA

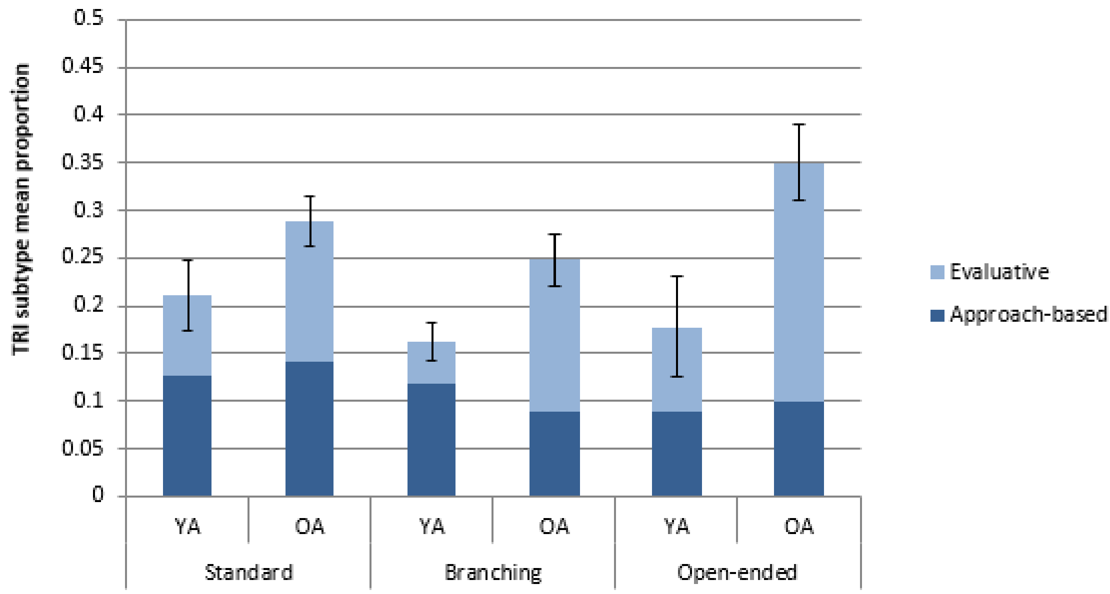


Figure 12. TRI subtypes as proportion of overall thought reports for the open-ended Study 1 participants, the non-branching probe Study 2 condition, and the branching probe Study 2 condition. Bars represent standard error for amount of TRI.

APPENDIX P

STUDIES 1-4 APPROACH-BASED TRI DATA

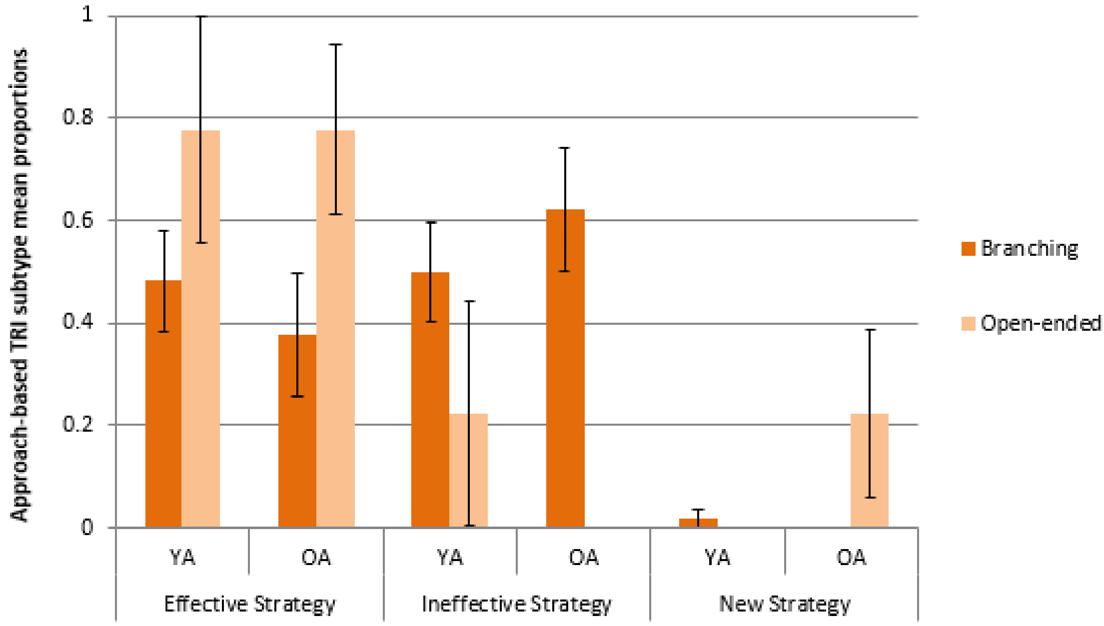


Figure 13. Approach-based TRI subtypes as proportion of overall reported approach-based TRI for the open-ended Study 1 participants and the branching probe Study 2 condition participants. Participants in the non-branching probe condition of Study 2 completed standard mind-wandering thought probes and did not answer questions about the specific types of approach-based TRI they experienced during the task. Bars represent standard errors.

APPENDIX Q

STUDIES 1-4 EVALUATIVE TRI DATA

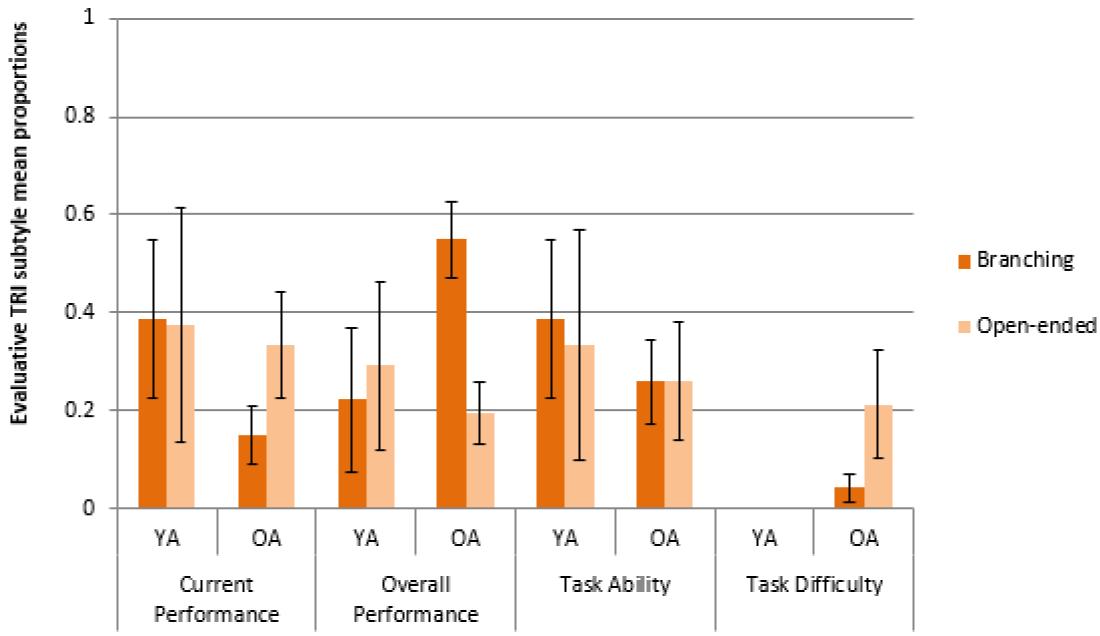


Figure 14. Evaluative TRI subtypes as proportion of overall reported evaluative TRI for the open-ended Study 1participants and the branching probe Study 2 condition participants. Participants in the non-branching probe condition of Study 2 completed standard mind-wandering thought probes and did not answer questions about the specific types of evaluative TRI they experienced during the task. Bars represent standard errors.

APPENDIX R

STUDY 5 DESCRIPTIVES BY AGE, CONDITION, AND BLOCK

		Younger Adult				*Older Adult*			
		JOL Only		*JOL + Probe*		*JOL Only*		*JOL + Probe*	
		Block 1	Block 2	Block 1	Block 2	Block 1	Block 2	Block 1	Block 2
TRI	*M*	0.054	0.025	0.074	0.041	0.075	0.03	0.169	0.153
	SE	0.016	0.013	0.016	0.014	0.016	0.014	0.016	0.014
Approach-Based TRI	*M*	0.004	0.001	0.01	0.003	0.008	0.001	0.017	0.009
	SE	0.004	0.001	0.004	0.002	0.004	0.001	0.004	0.002
Evaluative TRI	*M*	0.05	0.024	0.064	0.038	0.067	0.029	0.153	0.144
	SE	0.016	0.014	0.015	0.014	0.015	0.013	0.016	0.014
TUT	*M*	0.136	0.137	0.117	0.143	0.029	0.035	0.022	0.021
	SE	0.015	0.016	0.014	0.015	0.014	0.016	0.015	0.017
PA Recall	*M*	0.41	0.402	0.569	0.511	0.396	0.426	0.454	0.449
	SE	0.028	0.029	0.027	0.029	0.028	0.03	0.026	0.029
Mean JOL	*M*			53.78	51.266			55.798	51.332
	SE			2.655	2.804			2.655	2.804
JOL Simple Diff	*M*			0.032	-0.002			-0.104	-0.065
	SE			0.33	0.32			0.34	0.33
JOL Abs Diff	*M*			0.188	0.171			0.23	0.236
	SE			0.019	0.02			0.018	0.021
Gamma	*M*			0.266	0.282			0.256	0.224
	SE			0.029	0.026			0.03	0.025

Note. Only experimental participants were required to make JOLs in addition to responding to thought content probes.

APPENDIX S

STUDY 5 DESCRIPTIVES BY AGE, CONDITION, AND RELATEDNESS

		Younger Adult						Older Adult					
		JOL Only			JOL + Probe			JOL Only			JOL + Probe		
		Strongly Related	*Moderately Related*	*Unrelated*	*Strongly Related*	*Moderately Related*	*Unrelated*	*Strongly Related*	*Moderately Related*	*Unrelated*	*Strongly Related*	*Moderately Related*	*Unrelated*
TRI	*M*	.022	.019	.037	.031	.031	.053	.027	.029	.049	.104	.083	.135
	SE	.011	.01	.011	.009	.1	.012	.011	.01	.012	.01	.012	.011
Approach-Based TRI	*M*	.001	.001	.003	.004	.001	.007	.003	.002	.006	.006	.004	.017
	SE	.001	.001	.003	.002	.001	.003	.002	.002	.003	.002	.002	.003
Evaluative TRI	*M*	.021	.019	.034	.027	.031	.044	.024	.029	.043	.098	.08	.119
	SE	.01	.01	.011	.01	.01	.012	.009	.01	.011	.011	.01	.011
TUT	*M*	.075	.097	.087	.072	.109	.078	.022	.026	.016	.01	.016	.018
	SE	.009	.011	.011	.009	.01	.01	.008	.011	.011	.009	.011	.01
PA Recall	*M*	.497	.425	.293	.689	.587	.335	.524	.442	.272	.617	.501	.22
	SE	.029	.031	.03	.029	.031	.03	.029	.031	.03	.027	.032	.03
Mean JOL	*M*				.638	.528	.339				.675	.613	.229
	SE				.029	.031	.029				.032	.033	.031
JOL Simple Diff	*M*				.05	.005	-.004				-.057	-.112	-.079
	SE				.035	.037	.033				.035	.037	.033
JOL Absolute Diff	*M*				.166	.189	.171				.252	.262	.208
	SE				.02	.023	.022				.019	.023	.022
Gamma	*M*				.127	.111	.133				.049	-.003	.063
	SE				.027	.03	.035				.028	.031	.036
Effective Strategy	*M*	.574	.554	.481	.579	.559	.467	.585	.559	.488	.605	.578	.457
	SE	.016	.015	.02	.016	.015	.02	.016	.015	.02	.016	.015	.02

Note. Only experimental participants were required to make JOLs in addition to responding to thought content probes.