SMEEKENS, BRIDGET A., M.A. The Role of Working Memory Capacity and Mind Wandering in Creativity and Insight. (2013)
Directed by Dr. Michael J. Kane. 106 pp.

Conflicting theories suggest opposing predictions for the role of working memory capacity (WMC) and mind wandering in insight problem solving and creativity. The executive-control-benefit perspective suggests that insight problem solving and creativity would benefit from the effectively focused attention that high WMC enables. Focused attention should help guide a selective search of solution-relevant information in memory and help inhibit uncreative yet accessible ideas. In contrast, the executive-control-cost perspective suggests that unfocused attention would be beneficial to insight and creativity, as it should allow access to more loosely relevant concepts, remotely linked to commonplace ideas. By inserting incubation periods into two insight problems and two creativity tasks, my main goal was to test whether or not WMC and mind wandering during the incubation tasks predict post-incubation performance on insight and creativity problems. Yet a third possibility, however, is that individual differences in WMC predict *flexibility* in control, such that people with higher WMC better adjust attentional focus (i.e., narrowly or broadly) to fit the requirements of the task. For instance, there is less benefit to mind wandering during a stand-alone attention task than during the incubation period of an insight problem or creativity task, where task-unrelated thoughts could lead to progress toward the problem-solving or creative goal. Following up on previous research, I also explored the possibility, in both studies, that self-reported concentration during the attention-demanding tasks may moderate the relationship between WMC and mind wandering in the lab, as it does in daily-life activities. In a second experiment I included an openness measure and a need for cognition measure in order to assess the moderating role of intellectual motivation on WMC to predict success in insight and creativity. Overall, results suggest that WMC is beneficial for certain insight problems, but not for creativity, and whereas mind wandering is not helpful for creativity, it may in fact be harmful in some insight problems. In addition, concentration does not seem to interact with WMC to predict mind wandering in the lab like it does in daily life. Finally, although openness to experience predicted both TUTs and creativity, neither openness to experience nor need for cognition moderated the relationship between WMC and insight or creativity.

THE ROLE OF WORKING MEMORY CAPACITY AND MIND WANDERING IN CREATIVITY AND INSIGHT

by

Bridget A. Smeekens

A Thesis 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
Master of Arts

Greensboro
2013

Approved by

____
Committee Chair

APPROVAL PAGE

This thesis written by Bridget A. Smeekens has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ________________________________
Michael J. Kane

Committee Members ________________________________
Peter F. Delaney

____
Paul J. Silvia

8/15/2013____________________
Date of Acceptance by Committee

8/15/2013________________
Date of Final Oral Examination

ACKNOWLEDGMENTS

I thank my advisor, Dr. Michael J. Kane for his guidance and patience throughout this process. I also thank my committee members, Drs. Peter F. Delaney and Paul J. Silvia for their expertise and assistance with this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Variation in working memory capacity, in part, reflects individual differences in domain-general, attention control capabilities (e.g., Kane, Conway, Hambrick, & Engle, 2007) and positively predicts higher-order cognitive abilities such as reading comprehension (Daneman & Carpenter, 1980), reasoning (Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen & Christal, 1990; Salthouse, 1993), and analytical problem solving (Hambrick & Engle, 2003). According to an executive attention perspective, individual differences in domain-general executive attention allow for active maintenance of goal-relevant information in the face of response conflict and memory interference (Engle & Kane, 2004). Furthermore, this executive attention is thought to be important in analytical problem solving because it helps to maintain access to goals, plans, and other task-relevant information, and to inhibit task-irrelevant information (Hambrick & Engle, 2003), both of which help solvers progress towards the correct solution (Kaplan & Simon, 1990). Overall, then, the research seems to agree that having greater attention control is beneficial for many things. But, is it beneficial for everything? Recently, some work has suggested a role for executive control in creativity and insight, but the literature is mixed in terms of what that role is—some argue that executive control is helpful for insight and creativity, while others maintain that a *lack* of executive control is ideal.

In two studies I will attempt to determine the influence of executive control on creativity and insight by correlating individual differences in working memory capacity (WMC) with creativity and insight, and by borrowing a design from the incubation literature to measure mind wandering while taking a break from creativity and insight problems. I also measured concentration during the mind wandering tasks to see whether it moderates the relationship between WMC and mind wandering. Previous work has suggested that this relationship may change depending on how much a person is concentrating; therefore, it is possible that we may be underestimating the relationship between WMC and mind wandering by omitting concentration. Finally, exclusively in the second experiment I tested the hypothesis that intellectual motivation may interact with WMC to predict success in insight and creativity by adding measures of openness and need for cognition.

Defining Insight and Creativity

Insight problem solving is distinct from analytical problem solving, in that analytical problems benefit from the application of prior domain knowledge, whereas insight problems are initially hindered by prior knowledge. In analytical problem solving, previously learned information is necessary to orient one's approach to the problem (Newell & Simon, 1972); however, in insight problem solving, prior knowledge acts as a stumbling block when the solver becomes fixated on one particular method of solving or thinking that has previously been successful, but is not currently (Chronicle, MacGregor, & Ormerod, 2004; Kaplan & Simon, 1990; Ormerod, MacGregor, & Chronicle, 2002). In order to reach the correct solution, then, the solver must restructure the problem, or think about it in a new way (i.e., without the previous, unnecessary restraints; e.g., Ohlsson, 1992; Weisberg, 1995). In a sense, the difficulty in analytical problem solving comes from the size of the search space, or the number of possible answers that must be retrieved or worked through in order to find the right one (Hambrick & Engle, 2003; Newell & Simon, 1972), whereas the difficulty in insight problem solving comes from the placement of false restraints on the problem, which initially makes it impossible to solve (Ash & Wiley, 2006; Ohlsson, 1992).

Similarly, creativity, or the process of generating novel and useful interpretations of old items and ideas, is also thought to be hindered at times by previously acquired knowledge of norms and commonplace uses (e.g., Gilhooly, Fioratou, Anthony, & Wynn, 2007). It is only after these restraints are removed that one can think of innovative and creative ideas. For example, in the alternative uses task, a common laboratory task used to measure creative, or divergent, thinking (Guilford, 1967), subjects initially output "creative" uses for a common object that were simply retrieved from long term memory (Gilhooly, Fioratou, Anthony, & Wynn, 2007). While these mundane responses are easy to access, people who are able to inhibit thinking about and reporting common uses by switching to a more effective strategy perform better (Gilhooly et al., 2007). In a second experiment, Gilhooly et al. asked subjects to distinguish between alternative-use responses that they had retrieved from memory from those that they generated on the spot. Gilhooly et al. found that executive functioning (as measured by a letter fluency task) was significantly correlated with the number of novel responses that were generated during the experiment, but unrelated to the number of responses retrieved from memory.

Their proposed mechanism of inhibiting common uses for an everyday object implicates a role for executive attention, and therefore, WMC.

The Benefits of Executive Control to Insight and Creativity

Two opposing theories attempt to explain the mechanisms underlying insight. According to Gestalt theorists, insight results from an automatic and unconscious process (e.g., Knoblich, Ohlsson, Haider, & Rhenius, 1999; Ohlsson, 1992; Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995) whereby a sudden burst in neural activity allows solvers to realize relationships between components of the problem that were not previously perceived to be related (Jung-Beeman et al., 2004). Alternatively, others (e.g., Davidson, 1995; Kaplan & Simon, 1990; Weisberg & Alba, 1981) have suggested that insight results from a controlled search of potential problem representations. This search for problem representations in insight problems is argued to be similar to the controlled search that allows navigation through the problem space of an analytical problem when progressing towards a solution (Kaplan & Simon, 1990). Just as WMC is beneficial for this controlled search component in analytical problem solving (Hambrick & Engle, 2003), it follows that WMC would also play an important role in the controlled search within insight problems. Over the years, several researchers have attempted to elucidate this controversy by correlating performance on executive control tasks with insight problem solving rates and solution times (Ash & Wiley, 2006; Fleck, 2008; Gilhooly & Fioratou, 2009; Gilhooly & Murphy, 2005). The motivation for this approach is that if insight problems benefit from a controlled search process, then people with better attention control (who also have a higher WMC) will perform better on insight problems than those with worse attention control (and lower WMC). Alternatively, if a controlled search plays no role in solving insight problems, then abilities that require attention control, such as WMC, should not be correlated with insight problem solving rates and times.

In analyses comparing insight to analytical problems, WMC is typically found to be significantly correlated with both insight and analytical problem solving times, such that people with higher WMC solve the problems faster than those with lower WMC (but don't necessarily differ in solving rates; Fleck, 2008; Gilhooly & Fioratou, 2009; Gilhooly & Murphy, 2005). Notably, in these studies, WMC predicted insight and analytical problem solving equally well. As a follow up to this finding, Gilhooly and Fioratou found that visuo-spatial and verbal WMC significantly predicted insight problem solving when entered as a second step into a statistical model that already contained measures of attentional inhibition and switching. From this, they concluded that WMC's contribution to insight problem solving is through its storage capacity rather than its executive control mechanisms.

Ash and Wiley (2006) attempted to isolate the restructuring phase of insight problem solving (i.e., the point at which the solver rethinks the problem without the previous, self-imposed constraints), and WMC's role therein, by comparing two versions of six insight problems. One group of problems, termed Many Moves Available problems, allowed for many possible moves before hitting an impasse, or getting stuck (i.e., they presented a large problem space). In the other group, Few Moves Available problems, allowed for only a few possible moves before hitting impasse, minimizing the size of the problem search space before restructuring. WMC only predicted solving rates in Many Moves Available problems, and not Few Moves Available problems, suggesting WMC predicts insight problem solving when the initial search space, prior to restructuring, is large (i.e., when there are many wrong potential answers before hitting an impasse; Ash & Wiley, 2006), but that it is not important for restructuring. That is, restructuring does not benefit from controlled attention and is perhaps instead the result of some automatic processes. From these findings, they concluded that when the search space is large, a controlled search of the problem representations could be beneficial for restructuring. For example, high-WMC solvers, who are able to maintain unsuccessful problem attempts in memory, may notice a pattern among their failures that will help them move towards the correct answer. If problem solving success is facilitated by the ability to maintain failed attempts while searching a problem space for new solutions, then WMC, which allows for both the attention and memory capacity to do this, should be correlated with success in insight problem solving.

Consistent with Ash and Wiley's (2006) findings, a similar pattern can be seen in the nine-dot problem (Chein, Weisberg, Streeter, & Kwok, 2010). Subjects were presented with a picture of nine dots arranged in the shape of a (3×3) square and were instructed to use four straight lines to connect all of the dots without lifting their pen from the paper. Although it is not explicitly stated in the instructions of the problem, the box-shape that is created by the organizational pattern of the dots initially prevents subjects from drawing lines outside of the square (Newell & Simon, 1972; Weisberg & Alba, 1981). In order to reach the correct solution, the problem space must be restructured in a way that “allows” solvers to draw lines beyond the square. Chein et al. found that spatial WMC predicted this tendency to draw outside the box—a breakthrough that is necessary but not sufficient to solving. Although spatial WMC only predicted solving in a version of the task with hints, these results can be interpreted as support for the contribution of WMC via controlled search of numerous nine-dot problem representations. Together, then, all of these studies suggest that WMC plays a key role in insight problem solving, perhaps similar to its role in analytical problem solving, by allowing a more effective controlled search of the potential problem representations and in maintaining information relevant to the goal of the task.

WMC has also been found to positively predict performance on the Remote Associates Task—a verbal reasoning task that is commonly used to measure verbal and fluid abilities and sometimes used to measure insight and creativity. Typically, in the Remote Associates Task, subjects are presented with three words (e.g., *eight, skate, stick*) and must come up with a fourth word that creates a compound word with each of the prompts (e.g., “figure”). What makes a Remote Associates Task difficult is when subjects come up with a word that will make a compound word with one or two of the prompts, but not all of them. Kane et al. (2004) found positive correlations between six measures of WMC and a Remote Associates Task, ranging from small to medium (*r*s = .06 to .28; although in this experiment the Remote Associates Task was used as a measure of verbal reasoning, rather than creativity). Ricks, Turley-Ames, & Wiley (2007) also found a positive relationship between WMC and a modified version of the Remote Associates Task, but not for people with strong prior knowledge (e.g., baseball players). In two experiments, subjects completed WMC measure(s), a Remote Associates Task, and answered questions about their baseball knowledge. This Remote Associates Task was different from what has been used in other studies. In the version of the Remote Associates Task used for this study, the first two prompts (e.g., wild, dark) could be completed with a baseball-related word (e.g., pitch), however the third word would not fit that same pattern (e.g., sense). They found that in the neutral Remote Associates Task, WMC was positively correlated with proportion of correct solutions, regardless of baseball knowledge. Again, in the baseball-misleading Remote Associates Task, WMC was positively correlated with proportion of correct solutions; however, this time it was qualified by a significant interaction. People who had higher WMC and greater baseball knowledge actually performed worse than people with lesser baseball knowledge. Ricks et al. interpreted these results to mean that a higher WMC, or a more controlled attention, is beneficial for creative problem solving, except in the case of when prior knowledge activates an incorrect answer. When this happens, having more controlled attention is disadvantageous because it only helps the person focus more intensely on the wrong answer.

Although there is not much work that has directly assessed the relationship between WMC and divergent thinking, there is a growing literature that evaluates the relationship between general fluid intelligence and creativity. General fluid intelligence is the domain-free ability to reason through novel problems (Carroll, 1993; Horn & Cattell, 1967), usually analytical ones, such as in the Ravens Progressive Matrices test (Raven, Raven, & Court, 2003). In addition, general fluid intelligence correlates strongly with

WMC (typically around .70 - .80; see Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & Süβ, 2005); this association, according to an executive attention perspective, is thought in part to reflect a shared, domain-general executive attention ability (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999). Therefore, while I do not consider WMC and general fluid intelligence to be one in the same, the aspects of general fluid intelligence that have been found to be important in creativity (e.g., inhibition) might be those that are shared also with WMC's executive aspects.

Recent work finds that general fluid intelligence correlates positively with creativity in divergent thinking, alternative-uses, tasks (typically around .35 – .45 at the latent level; Gilhooly et al., 2007; Nusbaum & Silvia, 2011; Silvia, 2008). Nusbaum and Silvia's latent-variable analyses indicated, in their Study 1, that the relationship between general fluid intelligence and creativity was mediated by the executive process of switching between categories of uses (i.e., number of categories) but not clustering (i.e., number of words produced in each category). In their Study 2, half of the subjects were given a strategy to improve performance in the divergent thinking task, which involved disassembling the object and reassembling it. Instead of diminishing the effect of general fluid intelligence on creativity, instructing strategies amplified it, such that fluid intelligence better predicted performance on the divergent thinking task in the strategy group than in the control group (Nusbaum & Silvia, Study 2). Nusbaum and Silvia suggested that executive abilities are required to maintain a strategy in memory and use it effectively; therefore, those high in fluid intelligence will be able to do this successfully while those who are low in fluid intelligence will not. Gilhooly et al. (2007) suggests that this relationship stems from the prerequisite ability to inhibit commonplace ideas in the pursuit of inventing creative ones.

The Costs of Executive Control to Insight and Creativity

Several studies have found that people with lower WMC are more likely to mind-wander inappropriately (i.e., when the task at hand requires concentration) than are those with higher WMC (Kane et al., 2007; McVay & Kane, 2009, 2012a, 2012b; Unsworth & McMillan, 2013). Mind wandering is argued, by some theorists, to be automatically triggered by environmental and mental cues, and to represent a failure of the executive control system to block interference from thoughts unrelated to the ongoing task (McVay & Kane, 2010). In other words, mind wandering occurs when a person fails to inhibit off-task thoughts, pulling attention away from the current task, resulting in more errors in the task (e.g., McVay & Kane, 2009). Based on these negative consequences associated with attention control failure, one might predict that mind wandering would be harmful to performance in cognitively demanding tasks such as insight problem solving and creativity.

Alternatively, some have suggested that mind wandering may be beneficial for insight and creativity (Baars, 2010; Schooler et al., 2011; Smallwood & Schooler, 2006). Informal support for this idea can be found in countless anecdotes that illustrate fruitful mind wandering during an incubation period, a modern example of which includes chemist Kary Mullis solving the problem of how to amplify a target sequence of DNA (Mullis, 1998). In 1983, Mullis was driving through the mountains of California, pondering the principles of DNA replication when he was suddenly struck with insight at mile marker 46.58 on Highway 128. His realization of how to target a specific portion of DNA and then replicate it exponentially would eventually lead to the development of polymerase chain reaction (PCR) and secure his 1993 Nobel Prize in chemistry.

Indirect evidence from the incubation literature also supports a role for mind wandering in insight and creative problem solving. Over the years, dozens of studies have evaluated the effects of various incubation tasks (i.e., taking a break from an initial task to work on another task, before returning to the initial task) on insight problem solving and creativity. Some have found that an incubation period inserted into a problem is beneficial to subsequent creative problem solving (known as the "incubation effect"), while others have failed to find this effect. In an attempt to specify the factors that contribute to a successful incubation period, a recent meta-analysis categorized studies by problem type (e.g., creative, visual insight, or linguistic insight problems) and incubation task demands (e.g., rest, low cognitive demand, high cognitive demand), among other variables (Sio & Ormerod, 2009). Incubation periods filled with a low cognitive demand task were more beneficial for subsequent performance on creativity tasks than were high cognitive demand incubation tasks (although, there were not many data points to meta-analyze for the creativity problems; Sio & Ormerod). Likewise, incubation periods filled with a low cognitive load were optimal for both visual and verbal insight problems, compared to rest (Sio & Ormerod).

Based on these findings, Schooler et al. (2011) argued that an incubation task with low cognitive demands would allow abundant opportunities for mind wandering, whereas a task with high demands would not. In fact, in basic mind wandering research, low demand tasks yield higher mind-wandering rates than do high demand tasks (Smallwood, Nind, & O'Connor, 2009); therefore, these results suggest that mind wandering during a low-demand incubation period could be beneficial for creative problem solving. In addition, Smallwood and Schooler (2006) draw parallels between incubation processes and mind wandering, suggesting that the suddenness of insight "may sometimes occur because mind wandering addresses more remote goals (e.g., discerning the solution to a heretofore unsolved problem)" (Smallwood & Schooler, p. 956; see also Schooler et al., 2011). But, not all mind wandering is created equal—according to Smallwood, McSpadden, and Schooler (2008), tuning out (mind wandering with awareness) may be especially beneficial for insight and creativity, as compared to zoning out (mind wandering without awareness). They argue that even if a person has creative insight, if it occurred while the person was zoning out, it may go unnoticed. Instead, tuning out presumably provides the benefits of mind wandering while still allowing the person to recognize that a creative insight has occurred.

In the midst of data collection for my second experiment, Baird and colleagues (2012) published a paper apparently demonstrating that mind wandering during an incubation period was beneficial for creative problem solving. Their study compared different types of incubation tasks—an undemanding task, a demanding task, rest, and no break—during a divergent thinking task. Among other things, they hypothesized that subjects would mind wander more in the undemanding task compared to the demanding task, and would therefore show greater improvement in creativity from pre- to post-incubation. At the end of each incubation period, and before returning to working on the divergent thinking task, subjects reported on a 1–5 scale how frequently they were mind wandering during that incubation period. Although Baird and colleagues did not test (or, at least, report) the effect of mind wandering during the incubation period on improved creativity, they did find that the condition with the undemanding incubation task—the condition that reported the most mind wandering in a retrospective questionnaire—was the only condition to show significant improvement in creativity from pre- to post-incubation. In addition, a general propensity measure of daydreaming in daily life (Imaginal Process Inventory, Daydreaming Frequency subscale; Singer & Antrobus, 1972) positively correlated (*r* around .20) with post-incubation creativity scores, collapsing across all experimental groups. Thus, Baird and colleagues concluded that mind wandering is beneficial for creative problem solving.

Although Baird et al. are the only investigators to date to examine the influence of mind wandering on creativity, other research has implicated the benefits of decreased attentional control in creative success. Jarosz, Colflesh, and Wiley (2012) equated two groups of subjects on WMC using an operation span task, then left one group entirely sober and had the other group consume .88g/kg of body weight in vodka. An hour later, subjects completed another operation span task and then worked on a Remote Associates Task. While the sober group showed practice effects on the operation span task, the intoxicated group did not, suggesting decreased attentional control in those who consumed alcohol. In the Remote Associates Task, the intoxicated group solved more Remote Associates Task problems, did so faster, and was more likely to report that they had experienced insight. Jarosz et al. interpret these results to suggest that a decreased attentional control (i.e., a more diffuse attentional state) due to inebriation, as demonstrated by a lack of improvement in the operation span task, is beneficial for creative problem solving in the Remote Associates Task.

In another study, Aiello, Jarosz, Cushen, and Wiley (2012) found that taking a less-analytical approach to solving Remote Associates Task problems was significantly more successful than taking an analytical approach. Specifically, subjects who were given a "use your gut" instruction before completing Remote Associates Task problems performed better on the task compared to subjects who did not receive the "use your gut" instruction. The authors suggest that the "use your gut" instruction caused a reduced attentional control in subjects, leading to the conclusion that reduced attentional control is beneficial for solving creativity problems, such as the Remote Associates Task.

Other, tangentially-related, work also provides support for the argument that a *lack* of executive control is helpful for insightful and creative thinking. For example, White and Shah (2006, 2011) have found that adults with ADHD—a disorder with symptoms including reduced inhibitory control—perform better in divergent thinking tasks and have more real-world creative achievements than their non-ADHD counterparts. Another study looked at the relationship between time of day and problem solving and found that non-optimal time of day—which is also affiliated with reduced inhibitory control—was beneficial for solving insight problems but not analytical problems (Wieth & Zacks, 2011). Although neither of these studies measured attention control directly, they both draw conclusions that echo support for executive control's harmful influence on insight and creativity.

Flexibility in Executive Control

A third perspective, which combines these ideas, might suggest that WMC predicts flexibility in mind wandering, depending on context. That is, people high in WMC may be able to focus their attention (and reduce mind wandering) or unfocus their attention (and increase mind wandering) depending on what is most beneficial for the context they are in. Specifically, I predict that the correlation between WMC and mind wandering will be negative (around -.20; consistent with previous research) in the context of the stand-alone attention-demanding tasks, but closer to zero (or perhaps positive) in the context of the incubation tasks.

Such a flexibility hypothesis seems reasonable because the typical, negative association between WMC and mind wandering is not always seen. According to an executive control perspective, individual differences in WMC should be most pronounced when a task is difficult and requires concentration (Engle & Kane, 2004). If relevant goals and information are difficult to maintain while trying to complete a task, keeping that information in mind while keeping irrelevant information out will require greater executive control. So, if mind wandering is measured in the lab during a considerably easy task, which doesn't require much concentration to perform, there may be no relationship between WMC and mind wandering, or even a positive one (e.g., Baird, Smallwood, & Schooler, 2011; Levinson, Smallwood, & Davidson, 2012). A parallel example of this has also been seen outside of the lab in daily life research. Unlike typical lab tasks, activities in daily life vary substantially in terms of how much effort and concentration is required, from washing the dishes to mentally calculating how much tip to leave for the waitress. In their study, Kane and colleagues (2007) used palm pilots to spontaneously probe subjects about their thought content and context several times a day for one week. Their results showed a negative association between WMC and mind wandering, but only when concentration on the task was high. That is, when concentration was high, people with greater WMC mind-wandered less, but when concentration was low, people with greater WMC mind-wandered more. If concentration was not included in the model, however, WMC was unrelated to mind wandering. This finding suggests that, because concentration has never been taken into consideration in the laboratory, previous studies in the lab may have underestimated the relationship between working memory capacity and mind wandering.

The Present Work

Overall, the literature seems to be conflicted in its suggestions for how WMC and mind wandering might affect insight problem solving and creativity. On the one hand, the executive-control perspective predicts that higher WMC, and therefore a more focused attention, will benefit insight and creativity. On the other hand, the lack-of-executive-control perspective predicts that more mind wandering, and therefore a broader, unfocused attention, will be better for insight and creativity (c.f., White & Shah, 2006; 2011). Alternatively, a third prediction takes both of these perspectives into account, whereby WMC predicts flexibility in mind wandering, depending on context. In other words, people with higher WMC are better able to focus their attention and limit mind wandering when they need to, but they can also unfocus their attention and allow mind wandering when necessary or helpful, such as during an incubation period (Colflesh & Conway, 2007; Conway, Cowan, & Bunting, 2001).

In order to measure the impact of WMC and mind wandering on insight and creativity, I first created a composite WMC score for each person from two WMC tasks, operation span and symmetry span (Conway et al., 2005), and obtained baseline mind-wandering rates from a stand-alone, attention-demanding task. Two divergent thinking tasks and two insight problems were used to measure creativity and insight, respectively. An incubation period was inserted into the middle of each creativity and insight task, during which subjects responded to mind wandering probes regarding their current thoughts while completing an ongoing task that was comparable to the baseline measure.

The primary goal of this research is to examine the relationships among WMC, mind wandering state, insight, and creativity. Specifically, this study will assess the extent to which individual differences in WMC predict people's ability to control their attention, think creatively, and solve insight problems. I hypothesize that there will be an interaction between WMC and task context on mind wandering, such that people with lower WMC will mind wander more than people with higher WMC when there is no apparent reason to mind wander, but high spans will mind wander more than low spans during the incubation period when there is a potential reason to mind wander.

Previous work in the lab has shown a relationship between mind wandering and working memory capacity. Research into daily life has not found a raw correlation between mind wandering and working memory capacity; however, there is an apparent interaction between working memory capacity and concentration in predicting mind wandering in everyday life. This suggests that researchers could be underestimating the effect between mind wandering and working memory capacity in the lab by omitting concentration measures. In the present study, I measured concentration in order to test the interaction between it and working memory when predicting mind wandering. From an executive attention perspective, working memory should differentially predict mind wandering in conditions where concentration is high. That is, people with lower working memory should mind wander more than people with higher working memory when concentration is high.

CHAPTER II

EXPERIMENT 1

Assuming that an incubation period is beneficial to insight and creativity because it allows for mind wandering, I chose two tasks that I knew provided a supportive environment for off-task thought: the SART and the n-back. Previous work has consistently shown that subjects mind wander approximately 30–50% of the time while working on the SART (e.g., McVay & Kane, 2009; 2012a; McVay, Meier, Touron, & Kane, 2013). In addition, I chose the n-back because it has been used successfully as an incubation task before (Dijksterhuis & Meurs, 2006) and has been shown to induce comparable mind-wandering rates to my stand-alone measure, the SART (McVay, Meier, Touron, & Kane).

Methods

Subjects

Undergraduates enrolled in introductory psychology courses at the University of North Carolina at Greensboro (UNCG) earned partial credit toward a course requirement for their participation in each of two 120 min sessions. One hundred and seventy-three undergraduates completed the first session of the experiment; of those, 142 students returned to complete the second session as well. In order to minimize attrition, we gave the majority of the credits upon the completion of the second session. Unless stated otherwise, data analyses only included students who completed both sessions.

Working Memory Span Tasks

We assessed WMC with two “complex span” tasks that are commonly used to measure WMC in younger adults (Unsworth, Heitz, Schrock, & Engle, 2005). Both tasks—operation span and symmetry span—required subjects to perform a processing task while simultaneously remembering short lists of other, unrelated information in serial order. Each of the span tasks began with three practice blocks: the first block gave the subject an opportunity to practice the memory component of the task for 4 trials (two trials of set size two and two trials of set size three), the second block allowed for practice of the processing component of the task for 15 trials, and the third block combined both the memory component and the processing component in order to familiarize the subject with how the actual task would run for 8 trials. Following the third practice block, subjects were notified that they would begin the actual task. After the practice blocks, response deadlines for the processing task were used to ensure that the subjects did not rehearse the memory items during the processing component of the task. Response deadlines were calculated using the response times attained during the processing practice block (M + 2.5 SDs). If the subject exceeded the response deadline on any one trial, the task moved on to the next trial and counted that item as an error.

Operation span. In this task, subjects were asked to solve a math problem (e.g., (9 ÷ 3) + 2 = ?) and then evaluate the solution presented on the next screen (e.g., 6; true or false) while remembering a subsequently presented letter (see Unsworth et al., 2005). After three to seven of these equation–letter pairs, subjects saw the twelve possible letters (F, K, P, S, H, L, Q, T, J, N, R, and Y) and used the computer mouse to recall the letters in the order in which they had appeared. A total of 15 trials, ranging from set size three to seven (three of each set size), appeared in random order.

Symmetry span. In this task, subjects evaluated the vertical symmetry of a black-and-white 8×8 grid pattern while remembering the location of a subsequently presented red square within a 4×4 grid (see Kane et al., 2004). After two to five symmetry-square pairs, subjects saw an empty 4×4 grid and used the computer mouse to recall the locations of the red squares in the order in which they had appeared. A total of twelve trials, ranging from set size two to five (three of each set size), appeared in random order.

Mind Wandering Assessment: Ongoing Tasks and Thought Probes

Mind wandering was measured by randomly probing subjects about their thoughts during ongoing cognitive tasks. The Sustained Attention to Response Task (SART) served as a stand-alone measure and four different versions of the n-back served as the incubation tasks. Each took about 20 min to complete.

SART. This is a go/no-go task in which subjects viewed a sequence of words, presented one at a time, and had to decide whether each was an animal or a food (animals that are typically foods, such as chickens, were not included; see McVay & Kane, 2009). Each word was presented for 300 ms, followed by a 900 ms mask. If the word was an animal, subjects responded by pressing the space bar as quickly as they can (making up 88.9% of the 900 trials). If the word was a food, subjects withheld response, and therefore pressed no key (making up 11.1% of the 900 trials). Following 60% of the critical, no-go stimuli, subjects were presented with a thought probe, asking them to characterize the content of their thoughts in the moment before the thought probe appeared (see below for details).

N-back task. Subjects viewed a sequence of words, presented one at a time, and had to decide whether the current word is the same word that was presented two words back. If the word is the same (e.g., green, blue, *green*), subjects responded by pressing the space bar as quickly as they could (making up 25% of the 336 trials). If the word was not the same, subjects withheld response, and therefore pressed no key (making up 75% of the 336 trials). A subset of these non-target trials, called lure trials (making up 21% of the 336 trials), presented either 1-back matches (e.g., blue, green, *green*) or 3-back matches (e.g., green, blue, purple, *green*). Each word was presented for 500 ms, followed by a fixation cross for 2500 ms. While working on the task, subjects were periodically presented with a thought probe (just as in the SART), asking them to characterize the content of their thoughts immediately prior to the probe. Three thought probes were presented during each of the seven blocks, of which two thought probes were presented immediately after a target stimulus. Versions of this task functioned as an incubation-period task for the two insight problems and the two creativity tasks.

Specifically, there were four versions of the n-back task, presenting different stimuli: colors, countries, parts of the body, and musical instruments (see Appendix A for full list of items). I selected the words in each version based on categorical norms from Battig and Montague (1969) and Van Overschelde, Rawson, and Dunlosky (2004).

Mind wandering probes. Each probe screen asked, “What were you just thinking about?” Subjects’ responses were conveyed via key press and the on-screen choices included: (1) the current task; (2) my performance on the task; (3) off-task: tune out (knew it all along); (4) off-task: zone out (without knowing it). Simply put, "tuning out" is when a person is fully aware that he is mind-wandering, and zoning out is when a person doesn't realize that he is mind- wandering until something in the environment (like the thought probe) interrupts him or he catches himself. Specifically, subjects were told in the instructions that:

> During this experiment you will be asked at various points whether your attention is firmly directed towards the task, or alternatively you may be aware of other things than just the task.
>
> Occasionally you may find as you are reading the text that you begin thinking about something completely unrelated to what you are reading; this is what we refer to as "mind wandering." We believe there are two forms of mind wandering:
>
> TUNING OUT: Sometimes when your mind wanders, you are aware that your mind has drifted, but for whatever reason you still continue to read. This is what we refer to as "tuning out"—i.e., when your mind wanders and you know it all along.
>
> ZONING OUT: Other times when your mind wanders, you don't realize that your thoughts have drifted away from the text until you catch yourself. This is what we refer to as "zoning out"—i.e., when your mind wanders, but you don't realize this until you catch it. (Smallwood, McSpadden, & Schooler, 2007, p. 533)

This distinction between tuning out and zoning out *may* be important because it is possible that even if a creative insight has occurred while a person is zoning out, it may go unnoticed (Smallwood, McSpadden, & Schooler, 2008), and so only tune outs might be beneficial to creativity. It is not clear, however, whether this distinction between tuning out and zoning out is capturing anything more than a gradation from lessdistracted to more-distracted.[1] Therefore, responses of either "3" (tuning out) or "4" (zoning out) were both scored as a task-unrelated thought (i.e., TUT), or a mind wandering experience. Responses of "2", reflective of thoughts about one's own performance, were considered as "task related interference" (TRI; Smallwood, Riby, Heim, & Davies, 2006) that does not represent either fully on-task or off-task thought (McVay & Kane, 2009; 2012a; 2012b; McVay et al., 2013). My primary dependent variable of interest here, then, is overall TUT rate in each task; however, secondary analyses will distinguish between tuning out and zoning out.

Concentration probes. In addition to being asked about mind wandering, subjects were also asked about how much they were trying to concentrate on the current task (i.e., either the SART or the n-back). Immediately following each mind wandering probe, subjects saw a screen that said "I was trying to concentrate on the current task," along with a scale ranging from 1 (not at all) to 7 (very much).

Insight Problems

Subjects were given an insight problem to work on for 3 min. Regardless of whether or not they solved the problem during this pre-incubation period, subjects then switched to an incubation task (the n-back) for 20 min. If subjects had solved the problem during the pre-incubation period, then the task ended at the conclusion of the incubation

[1] Schooler and colleagues have suggested in numerous studies (e.g., Schooler, Reichle, & Halpern, 2004; Smallwood, McSpadden, & Schooler, 2007) that there is a qualitative distinction between tuning out and zoning out; however, research has only established a single dissociation, where performance in an ongoing task is worse while zoning out than tuning out. Therefore, it is not clear to me that this is a valid distinction (i.e., that it's different from simply being less versus more off task) and so I will proceed with caution while using these terms.

period. If subjects had not solved the problem during the pre-incubation period, then they returned to the original insight problem for an additional 5 min. The two insight problems used were the coins problem (Ash & Wiley, 2006; Ormerod, Macgregor, & Chronicle, 2002) and the pigpen problem, (Gilhooly & Murphy, 2005; Schooler, Ohlsson, & Brooks, 1993; for illustrations of both problems, see Figure 1A and 1B) which were chosen based on a pilot study. Both problems are considered pure insight problems using Weisberg's (1995) taxonomy, which means that they require a restructuring of the original problem representation in order to be solved. In the coins problem, subjects had to reorganize an arrangement of eight coins so that each coin touched exactly three other coins. In order to reach the correct solution, subjects must restructure their two-dimensional perspective of the problem and stack two of the coins on top of the remaining coins. Likewise, in the pigpen problem, subjects were asked to add two square enclosures to a picture of nine pigs in a pen so that each pig is in a pen by itself. In order to reach the correct solution, subjects must restructure how they think of a square and place one square in side of a diamond (i.e., a square turned 45 degrees).

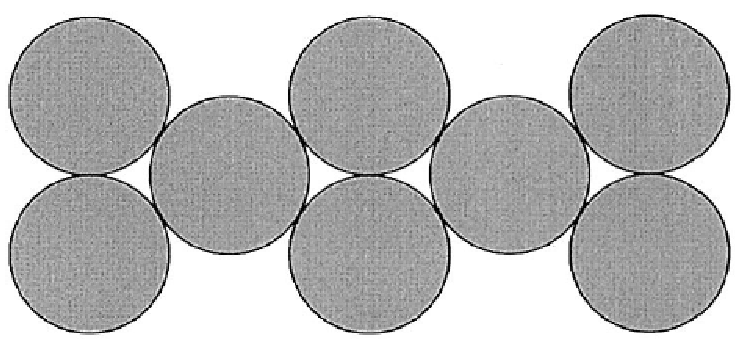

Figure 1A. Coin Insight Problem (Ormerod et al., 2002). There are 8 coins in this picture. Move 2 coins so that each coin touches exactly 3 other coins.

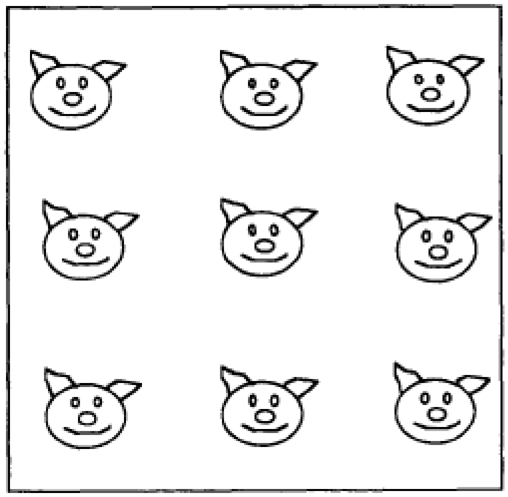

Figure 1B. Pigpen Insight Problem (Schooler et al., 1993). Nine pigs are kept in a square pen. Build two more square enclosures that would put each pig in a pen by itself.

Subjects saw the problems presented on a computer screen, one at a time, for a total of eight minutes each (if they did not solve the problem prior to incubation). During this time, subjects were given a blank piece of paper to work out their thoughts. If subjects thought they had the correct answer before time had expired, they pressed a key on the computer keyboard and a screen appeared instructing them to show their answer to the experimenter. The experimenter then checked the answer and recorded the accuracy of the response by pressing Y if the answer was correct and N if the answer was incorrect. If the answer was correct, subjects immediately moved on to the next part of the task; if the answer was not correct, the problem re-appeared (indicating to subjects that their previous response was incorrect) and subjects were given a new piece of scrap paper and allowed to work on the problem until time was up. I allowed this kind of feedback (and no more) to ensure that subjects did not give up after falsely believing they had solved the problem correctly.

Analytical Problems

Because the placement of false constraints plays such an important role in insight problems, the analytical problems were used as filler items to prevent subjects from noticing a pattern in the insight problems. My pilot study, which evaluated 11 different insight problems, suggested that once subjects realized that there was a "trick" to solving the problems, subsequent problems were more likely to be solved and were reported to be easier. Here, subjects tried to solve the crime problem (Schooler et al., 1993) and one free-response item adapted from the Ravens Advanced Progressive Matrices (i.e., item 23; Raven, Raven, & Court, 2003), presented on the computer screen, each for 3 min. In the crime problem, subjects were presented with statements from four suspects, only one of which is veridical. The problem stated "The police were convinced that Andy, Bill, Carl, or Dave had committed a crime. Each of the suspects made a statement, but only one of the statements was true. Andy said, 'I didn't do it.' Bill said, 'Andy is lying.' Carl said, 'Bill is lying.' Dave said, 'Bill did it.' Who is telling the truth and who committed the crime?" Given this information, subjects were instructed to write down on a piece of paper who was telling the truth and who committed the crime. In the Ravens problem, subjects were presented with a 3×3 matrix of abstract designs with the bottom right picture missing. Using the patterns throughout the rest of the matrix, the subject had to induce the rule that governed the progression of the figures left-to-right and top-to-bottom and then draw on a piece of paper what the missing piece of the matrix should look like. The procedure for these problems was identical to the insight problems (but without an incubation period).

Creativity (Divergent Thinking) Tasks

Two versions of the “alternative uses” task of divergent thinking (e.g., Guilford, 1967) were used to measure individual differences in creativity. In both versions, subjects were asked to generate as many creative uses for an everyday object as they can, for a total of 10 min. After 5 min on the task, subjects switched to the incubation task (the n-back) for approximately 20 min, and then switched back to the same alternative uses task for another 5 min. In one version of the task, subjects generated creative uses for a brick, and in the other version, a knife. Following recent work with divergent thinking tasks which strive to assess *creative* thinking (e.g., Silvia et al., 2008), my instructions emphasized that subjects should list creative, clever, original, unusual, and uncommon uses that are unlike any uses that they had seen or heard of before. After time was up, subjects were presented on-screen with a list of their responses and were asked to select what they thought were their two best answers. Subjects were asked to do this at three different times: once for their pre-incubation responses, once for their post-incubation responses, and once for all of their responses combined (pre- and post-incubation). Asking subjects for their top two responses allowed me to use top-two scoring, in addition to averaging across all of the responses a subject came up with in order to achieve an average creativity measure. More information about scoring will be provided in the section below.

General Procedure

Subjects individually completed two 120 minute sessions. All tasks were presented on computers. In the first session, subjects completed the SART, one of the alternative uses tasks, symmetry span, and one of the insight problems. In the second session, they completed a demographic survey, the crime problem, operation span, the other insight problem, the other version of the alternative uses task, and the Ravens problem. For both sessions, the order of the tasks was fixed for all subjects and administered in the aforementioned order. The insight problems and alternative uses tasks were counterbalanced so that half of subjects saw the coins and brick problem in session one and half saw the pigpen and knife problem in session one.

Scoring

Working memory tasks. Both the operation span task and symmetry span task were scored using partial credit load scoring (Conway et al., 2005), in which the total number of items recalled in correct serial position is summed across the task. These span scores were then individually converted into z-scores, based on our lab database of N = 3393 UNCG students, and then averaged to create a WMC composite.

Mind wandering tasks. Off-task thought reports (i.e., tuning out and zoning out) were categorized as mind wandering. The total rate of mind wandering responses to thought probes was calculated for each incubation n-back task and the SART. Initially, tuning out and zoning out were grouped together as task-unrelated thoughts (TUTs), but supplementary analyses, which are secondary to this study, will contrast tuning out and zoning out as two specific types of mind wandering.

Insight problems. Insight problems were scored as either 0 (incorrect) or 1 (correct).[2]

Creativity tasks. Three raters (the author and two other psychology graduate students who study creativity) scored each subject's individual responses on a scale of 1–5. For rating purposes, the judges were told to view creative ideas as having three facets: they are uncommon, they are remotely linked to everyday objects and ideas, and they are often clever. The raters saw all of the responses, from all subjects, from both experiment one and two, presented in alphabetical order in a spreadsheet, independent of any identifying information (see Table 1 for inter-rater correlations).

Table 1

Inter-rater Reliability for Divergent Thinking Tasks Across Experiments

	Brick Task		Knife Task	
	Rater 1	Rater 2	Rater 1	Rater 2
Brick Task				
Rater 1	1.000			
Rater 2	0.408	1.000		
Rater 3	0.454	0.417		
Knife Task				
Rater 1			1.000	
Rater 2			0.543	1.000
Rater 3			0.372	0.395

Note. Brick α = 0.688, Knife α = 0.689

[2] Since experimenters logged the accuracy of the insight responses by hand into the computer, there was opportunity for human error. To reduce this error, the responses in the output were compared to the hard copy of the subjects' answers (i.e., the drawings) by four different research assistants, entirely independent from one another. Then, those were compared to one another to confirm accuracy. Any discrepancies were decided upon by the author.

After all of the rating was complete, creativity was calculated for each subject in three different ways: (1) an *average* score was calculated by taking an average of all the ratings of one particular subject's responses; (2) a *top-two* score was calculated by averaging the ratings for the two responses that a subject indicated as the two best; (3) a *max-two* score was calculated by averaging the two responses that were given the highest mean ratings by the raters. For all of these creativity scores, ratings were first averaged across raters for each response, and then across all of the responses for each person. Previous work has found that average scoring is a bit more reliable than top-two scoring, but top-two scoring has greater validity than average scoring, insofar as it better predicts creative personalities (Silvia et al., 2008). I added max-two scoring in order to see whether it related to WMC and mind wandering differently than top-two scoring. In top-two scoring, subjects must be able to evaluate the creativity of their responses after generating them, but in max-two scoring, the evaluation of creativity is done only by the raters.[3]

Results

For all analyses, I report null hypothesis significance tests with an alpha of .05 and Cohen's d as a measure of effect size.

[3] Because RAs coded all of the subjects' responses for whether or not the subject had indicated it was one of their best two responses, there was once again a potential for human error. In order to minimize this, responses were coded by four different people, entirely independent from one another. Then, those were compared to one another to confirm accuracy. Any discrepancies were decided upon by the author.

Subjects

Data from subjects were omitted from analyses for scoring less than 85% on the processing component of either span task (18 people) and exceptionally poor performance on the tasks with embedded thought probes (SART and/or n-back tasks; 4 people); "poor performance" outliers were determined by collapsing non-target accuracy across n-back tasks and excluding anyone who had an accuracy of ≤ 75% in either the collapsed n-back or on the non-target trials of the stand-alone SART (mean accuracy and reaction times can be found in Appendix B). A total of 120 people were included for analyses, ranging in age from 18-29 years (M = 19.12, 65.8% female). By self-report, the final sample's racial composition was 58.3% White, 28.3% Black, 5.8% Asian, 5.0% Multiracial, and 2.5% Other. Additionally, regarding self-reported ethnicity, 5% self-identified as Latino/Hispanic.

Primary Analyses

Analyses that are central to my hypotheses will be included in the current section, while secondary questions and exploratory analyses will be presented in the next section. Descriptive statistics for WMC and TUT measures can be found in Table 2 and correlations among these variables can be found in Table 3. Standardized WMC scores were normally distributed with a mean close to zero. In addition, the individual WMC tasks (operation span and symmetry span) correlated fairly well, which allowed me to collapse across the two tasks to create composite WMC z-scores. Comparing TUT rates across tasks, people appeared to consistently mind wander about 40% of the time across tasks, regardless of whether the task stood alone (i.e., the SART) or was inserted as an incubation period for another task (i.e., n-back). A one-way analysis of variance (ANOVA) indicated no statistical differences in average TUT rates across any of the five different tasks, $F(4, 592) = 0.713$, $p = .583$, $MSE = 0.059$, $\eta_p^2 = .005$). Furthermore, TUT rates correlated significantly across all pairwise tasks, *r*s = .24 to .65 (see Table 3), suggesting that mind wandering is also consistent across people.

Table 2

Experiment 1 Descriptive Statistics for Working Memory Capacity Tasks and Mind Wandering

	N	Mean	SD	Min	Max	Skew	Kurtosis
Sspan Z	120	-0.062	0.975	-2.742	1.832	-0.396	-0.094
Ospan Z	120	0.061	0.949	-3.279	1.546	-0.941	1.258
WMC Z	120	-0.001	0.782	-1.637	1.589	-0.253	-0.613
SART TO	119	0.237	0.143	0	0.650	0.481	-0.135
SART ZO	119	0.130	0.113	0	0.483	0.963	0.604
SART TUT	119	0.367	0.196	0.033	0.800	0.203	-1.072
Coins TO	120	0.266	0.216	0	0.952	1.116	1.147
Coins ZO	120	0.146	0.153	0	0.667	1.184	1.047
Coins TUT	120	0.413	0.260	0	1	0.272	-0.750
PigPen TO	119	0.246	0.209	0	1	1.296	2.390
PigPen ZO	119	0.153	0.167	0	0.857	1.833	4.206
PigPen TUT	119	0.399	0.252	0	1	0.396	-0.489
Brick TO	120	0.246	0.193	0	1	1.402	2.922
Brick ZO	120	0.151	0.163	0	0.857	1.471	2.540
Brick TUT	120	0.397	0.236	0	1	0.438	-0.462
Knife TO	119	0.257	0.224	0	1	1.180	1.210
Knife ZO	119	0.156	0.159	0	0.905	1.669	4.059
Knife TUT	119	0.413	0.261	0	1	0.432	-0.575

Note. Z = Z scores based on database of over 3000 people; WMC = working memory capacity; SART = sustained attention response task; TO = tune outs; ZO = zone outs; TUT = task-unrelated thoughts

Table 3

Experiment 1 Correlations Among Working Memory Capacity Tasks and Mind Wandering

	1	2	3	4	5	6	7	8	9	10
1. Sspan Z	1.00									
2. Ospan Z	.32**	1.00								
3. WMC Z	.82**	.81**	1.00							
4. SART TOs	-.03	-.05	-.05	1.00						
5. SART ZOs	-.03	.04	.01	.16	1.00					
6. SART TUTs	-.04	-.01	-.03	.82**	.70**	1.00				
7. Coins TOs	-.20*	-.10	-.19*	.27**	.16	.29**	1.00			
8. Coins ZOs	-.09	-.12	-.13	.05	.41**	.27**	-.04	1.00		
9. Coins TUTs	-.22*	-.15	-.23*	.25**	.38**	.40**	.81**	.56**	1.00	
10. Pigpen TOs	-.15	-.12	-.17	.32**	.17	.33**	.39**	.01	.33**	1.00
11. Pigpen ZOs	-.04	-.08	-.07	-.01	.48**	.27**	.02	.52**	.32**	-.11
12. Pigpen TUTs	-.15	-.15	-.19*	.26**	.46**	.45**	.33**	.35**	.48**	.75**
13. Brick TOs	-.06	-.01	-.05	.26**	.06	.23*	.61**	-.05	.48**	.40**
14. Brick ZOs	.02	-.06	-.03	.04	.41**	.26**	.02	.59**	.37**	.02
15. Brick TUTs	-.04	-.05	-.06	.24**	.33**	.36**	.52**	.37**	.65**	.34**
16. Knife TOs	-.10	-.15	-.15	.37**	.16	.36**	.23*	.19*	.30**	.55**
17. Knife ZOs	.10	.05	.09	-.02	.50**	.28**	-.05	.48**	.24**	-.08
18. Knife TUTs	-.03	-.10	-.08	.31**	.44**	.48**	.17	.45**	.41**	.42**
19. N-back TOs	-.18	-.17	-.21*	.41**	.22*	.42**	.76**	.05	.66**	.80**
20. N-back ZOs	-.01	-.07	-.04	.03	.55**	.33**	-.01	.79**	.46**	-.05
21. N-back TUTs	-.14	-.17	-.19*	.34**	.53**	.54**	.58**	.56**	.81**	.59**

Table 3 (continued)

	11	12	13	14	15	16	17	18	19	20
1. Sspan Z										
2. Ospan Z										
3. WMC Z										
4. SART TOs										
5. SART ZOs										
6. SART TUTs										
7. Coins TOs										
8. Coins ZOs										
9. Coins TUTs										
10. Pigpen TOs										
11. Pigpen ZOs	1.00									
12. Pigpen TUTs	.57**	1.00								
13. Brick TOs	-.05	.30**	1.00							
14. Brick ZOs	.53**	.37**	-.13	1.00						
15. Brick TUTs	.32**	.50**	.73**	.59**	1.00					
16. Knife TOs	.03	.47**	.06	.00	.05	1.00				
17. Knife ZOs	.61**	.34**	-.05	.53**	.33**	-.11	1.00			
18. Knife TUTs	.40**	.61**	.02	.32**	.24**	.80**	.52**	1.00		
19. N-back TOs	-.04	.64**	.69**	-.02	.55**	.65**	-.09	.49**	1.00	
20. N-back ZOs	.83**	.51**	-.08	.82**	.50**	.05	.81**	.54**	-.03	1.00
21. N-back TUTs	.52**	.83**	.48**	.53**	.75**	.53**	.47**	.73**	.75**	.64**

Note. Z = Z scores based on database of over 3000 people; WMC = working memory capacity; SART = sustained attention response task; TO = tune outs; ZO = zone outs; TUT = task-unrelated thoughts.

* $p < .05$; ** $p < .01$

In the coins problem, there were 9 pre-incubation solvers (7.5%), and only 7 post-incubation solvers (5.8%), leaving 104 people who never solved the problem (86.7%). The pigpen problem also had 9 pre-incubation solvers (7.5%), 14 post-incubation solvers (11.7%), 96 people who never solved (80.0%), and one person who was excluded from pigpen analyses because he had seen the problem before (0.8%). Across all primary analyses involving either insight problem, people who solved pre-incubation were excluded.

Inter-rater reliability for divergent-thinking scores across experiments was quite good for both brick ($\alpha = 0.69$) and knife ($\alpha = 0.69$; see Table 1; c.f., Silvia et al., 2008) tasks. In Experiment 1, overall creativity ratings (collapsed across both pre- and post-incubation) in the brick task was correlated with overall creativity in the knife task, regardless of using average scoring, $r(117) = .58$, $p < .001$, top-two scoring, $r(117) = .38$, $p < .001$, or max-two scoring, $r(117) = .59$, $p < .001$. Therefore, for the creativity analyses, scores will be averaged across brick and knife in order to create one creativity measure.

Is executive control helpful for insight and creativity? Based on the perspective that executive control is helpful for insight and creativity, WMC should correlate positively with insight and creativity. Regarding insight, people who solve the coins problem after incubation did not have a significantly higher WMC z-score (M = -.06, SD = 1.09) than did those who never solved (M = .02, SD = 0.77), $t(109) = 0.24$, $p = .81$, 95% CI [-0.54, 0.69], d = 0.08; however, people who solved the pigpen problem after incubation did have significantly higher WMC z-scores (M = .60, SD = 0.77) than did those who never solved (*M* = -.11, *SD* = 0.76), *t*(108) = -3.28, *p* = .001, 95% CI [-1.14, -0.28], *d* = -0.93. In addition, there was no correlation between WMC and post-incubation creativity scores on the divergent thinking tasks for average, *r*(114) = .01, *p* = .90, top-two, *r*(114) = .07, *p* = .47, or max-two, *r*(114) = .03, *p* = .72 scoring (see Figure 2).

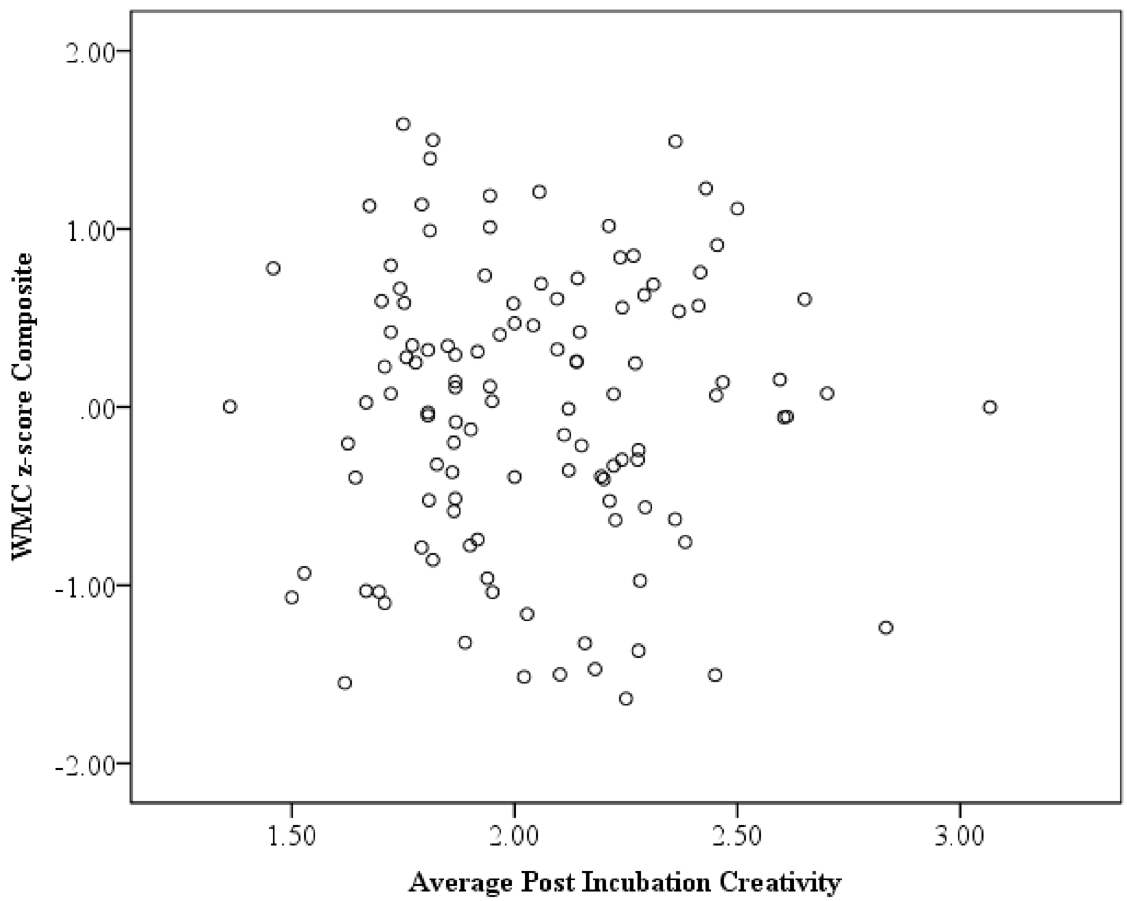

Figure 2. Experiment 1 Correlation Scatter between WMC and Creativity (Average Scoring).

Is lack of executive control helpful for insight and creativity? Based on the perspective that a *lack* of executive control is helpful for insight and creativity, TUT rates should positively correlate with post-incubation insight solution rates and creativity.

What I found, however, was that subjects who solved the coins insight problem after incubation did not mind wander significantly more than did those who never solved; if anything, they mind wandered somewhat less (*M* TUT rates = .35 [*SD* = 0.16] and .41 [*SD* = 0.26], respectively), $t(109) = 0.60$, $p = .55$, 95% CI [-0.14, 0.26], $d = 0.28$. Similarly, subjects who solved the pigpen insight problem after incubation tended to have fewer TUTs during the pigpen incubation period (*M* TUT rate = .30, *SD* = 0.29) than did subjects who never solved the pigpen problem (*M* TUT rate = .42, *SD* = 0.24), but not significantly so, $t(108) = 1.71$, $p = .09$, 95% CI [-0.02, 0.26], $d = 0.46$. Focusing on the incubation period during the divergent-thinking tasks, I found no relation between mind wandering and post-incubation creativity for average, $r(114) = -.08$, $p = .42$, top-two, $r(114) = -.10$, $p = .29$, or max-two, $r(114) = .01$, $p = .92$, scoring methods (see Figure 3).

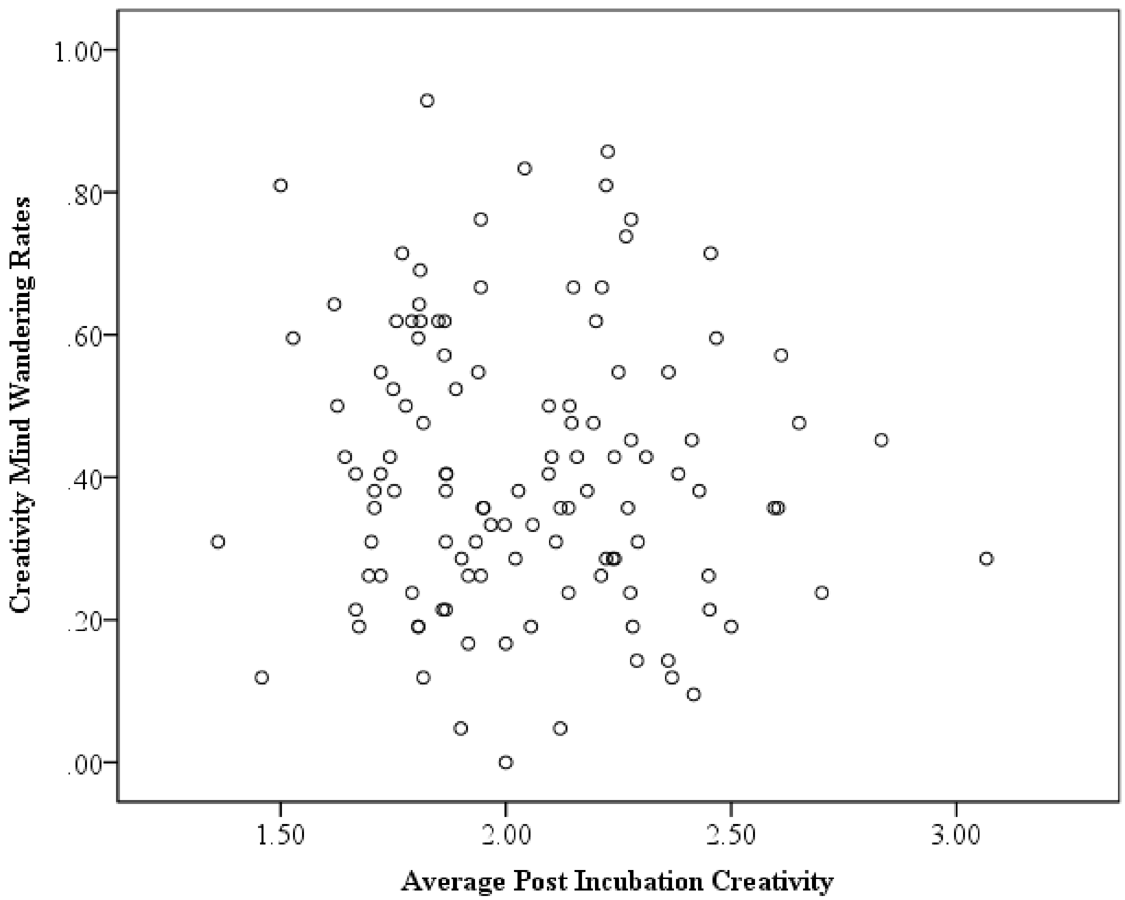

Figure 3. Experiment 1 Correlation Scatter between TUT Rates and Creativity (Average Scoring).

Although WMC and TUT rates related inconsistently to insight and creativity, people who are high in WMC *and* TUT rates may be more insightful or creative. To test whether or not this interaction between WMC and TUT rates could predict insight, I ran two separate logistic regressions—one for coins and one for pigpen (see Table 4). In the logistic regression with post-incubation coins solving as the outcome measure, I first entered WMC and coins TUT rates at Step 1 and added the WMC x TUT interaction at Step 2, but the interaction was not significant. A similar logistic regression was run for the pigpen problem, with post-incubation pigpen solving as the outcome measure and pigpen TUT rates replacing coins TUT rates. This time, the interaction was statistically significant, but in the negative direction, suggesting that post-incubation pigpen solvers had higher WMC and *lower* pigpen TUT rates.

Table 4

Experiment 1 WMC x TUT Logistic Regression on Insight Problem Solving

						95% CI for Exp(B)	
	B	*SE*	*Wald*	*p*	*Exp(B)*	*Lower*	*Upper*
Coins Predictors							
Block 1							
WMC z-score	-0.147	0.396	0.137	.711	0.863	0.397	1.877
Coins TUT Rate	-0.283	0.430	0.435	.510	0.753	0.324	1.749
Constant	-2.734	0.404	45.802	.000	0.065		
Block 2							
WMC x TUT	-0.428	0.426	1.010	.315	0.652	0.283	1.502
Pigpen Predictors							
Block 1							
WMC z-score	1.058	0.386	7.497	.006	2.880	1.351	6.141
Pigpen TUT Rate	-0.379	0.342	1.227	.685	0.685	0.350	1.339
Constant	-2.357	0.402	34.397	.000	0.095		
Block 2							
WMC x TUT	-0.905	0.387	5.451	.020	0.405	0.189	0.865

Note. WMC = working memory capacity; TUT = task-unrelated thoughts; Coins N = 111, Pigpen N = 110; all predictors are z-scores (for centering purposes).

For the creativity analysis, I used average post-incubation creativity scoring as the outcome measure in a hierarchical linear regression with WMC and creativity TUT rates entered in at Step 1 and the interaction of the two (WMC x TUT) entered in at Step 2, but the interaction effect was not significant (see Table 5).

Table 5

Experiment 1 WMC x TUT Hierarchical Linear Regression on Creativity

Predictors	*B*	*SE*	*β*	*t*	*p*	*95% CI for B* Lower	*Upper*
Block 1 (Δ R^2 = .006)							
WMC z-score	0.002	0.029	0.006	0.068	.946	-0.056	0.059
Creativity TUT Rate	-0.023	0.029	-0.075	-0.792	.430	-0.080	0.034
Constant	2.044	0.029		71.459	.000	1.987	2.101
Block 2 (Δ R^2 = .010)							
WMC x TUT	0.034	0.031	0.102	1.086	.280	-0.028	0.095

Note. WMC = working memory capacity; TUT = task-unrelated thoughts; all predictors are z-scores (for centering purposes).

Does the correlation between WMC and TUTs vary by context? A third hypothesis was that the relation between WMC and mind-wandering rates would change depending on the task context, demonstrating flexibility in executive control. I predicted that mind-wandering rates would be lower for people with high WMC than for people with low WMC when there was no experimentally-induced reason to mind-wander (i.e., during the SART), but greater for those with high WMC compared to those with low WMC when there was an experimentally-induced reason to mind-wander (i.e., during the n-back incubation period). To test this, I statistically compared the correlation (Steiger, 1980) between WMC and TUT rate during the SART (r = -.03) to the correlation between WMC and average TUT rate across the four incubation tasks (r = -.19). These two correlations did not differ significantly, $t(114) = 1.81$, $p = .07$; if anything, the pattern trended in the opposite direction of what I had predicted, with high spans mind wandering at about the same rate as low spans during the stand-alone SART, but less than low spans during incubation tasks.

Does concentration moderate the relationship between WMC and TUTs? Based on previous research in daily life (Kane et al., 2007), I predicted that WMC would significantly moderate the relationship between TUT rate and concentration, such that people with lower WMC should TUT more than people with higher WMC when self-reported concentration was high. For this analysis, I collapsed across all SART and n-back tasks, measuring mind wandering on any given trial as either "0" (on task) or "1" (off task). Because the outcome measure was binary and the concentration responses (Level 1 data) were nested within subjects (Level 2 data), I used hierarchical linear modeling (HLM) for binary outcomes with a Bernoulli distribution to test my hypothesis (Raudenbush & Bryk, 2002). The Level 1, within-subject variable—concentration—was group centered, and the Level 2, between-subject variable—WMC—was grand-mean centered. My model included fixed effects for concentration at Level 1 and fixed effects for WMC on the intercept and on the concentration slope (a WMC × concentration interaction) at Level 2. I also added a random effect for the intercept.

Although I did find that concentration negatively predicted TUTs (β_{10} = -1.04, $t(13975) = -18.91$, $p < 0.001$; Table 6), WMC did not significantly predict TUTs (β_{01} = -0.22, $t(118) = -1.62$, $p = 0.11$), and concentration did not moderate the association between WMC and TUTs (β_{11} = 0.02, $t(13975) = 0.34$, $p = 0.74$; see Figure 4). Finally, a significant person-level random effect on the intercept suggests that there were individual differences in TUT rates beyond what WMC and concentration can account for.

Table 6

Experiment 1 WMC x Concentration Hierarchical Linear Model on Mind Wandering

Fixed Effect	Coefficient	*SE*	*T-ratio*	*d.f.*	*p*
Intercept, β_{00}	-0.036	0.122	-0.292	118	.771
WMC, β_{01}	-0.218	0.135	-1.619	118	.108
Concentration, β_{10}	-1.044	0.055	-18.908	13975	.000
WMC, β_{11}	0.022	0.067	0.336	13975	.737

Random Effect	χ^2	*SD*	*Variance Component*	*d.f.*	*p*
Intercept, r_0	2599.634	1.325	1.755	118	.000

Note. WMC = working memory capacity; calculated by averaging across operation span and symmetry span z-scores. In this analysis, the main effects of WMC and concentration are represented by β_{01} and β_{10}, respectively, and the interaction between the two is represented by β_{11}.

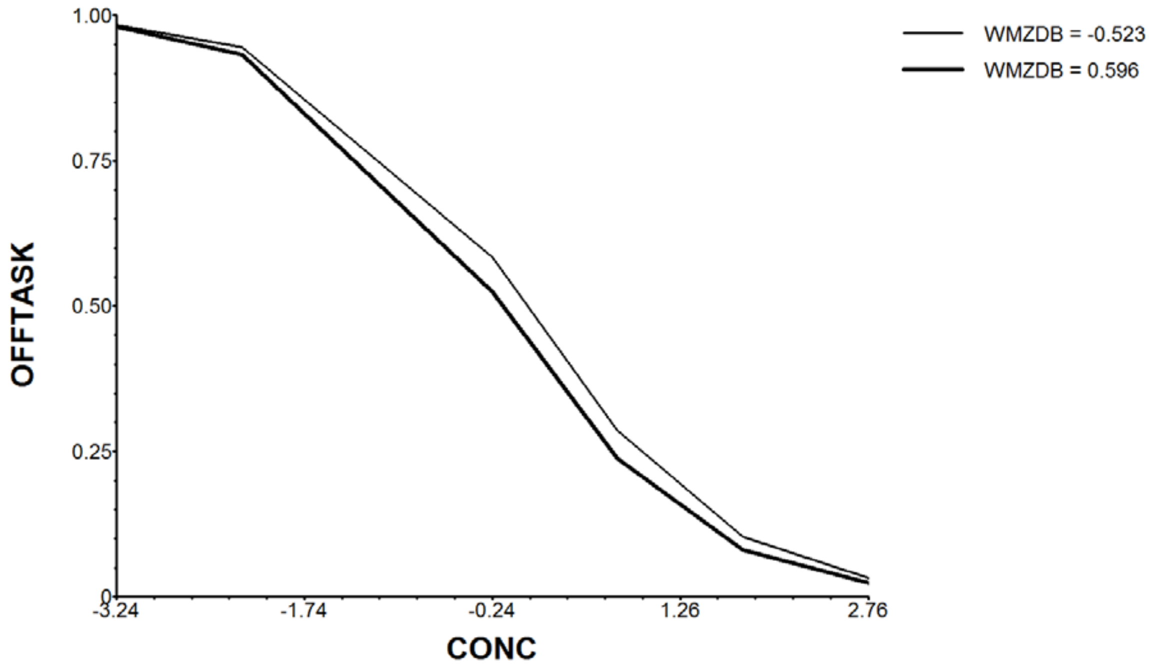

Figure 4. The Relation between Mind Wandering and Concentration for Subjects with High and Low WMC in Experiment 1. The lines represent the means of the within-person slopes for subjects in the top quartile of WMC (thicker line) and the bottom quartile of WMC (thinner line). Mind wandering is on the y-axis, where 0 = on-task and 1 = off-task. Concentration is group-centered on the x-axis, where negative values indicate less concentration and positive values indicate more concentration.

Secondary Analyses

Although I was primarily interested in how TUTs related to post-incubation insight and creativity, I also ran similar models distinguishing tune outs from zone outs as predictors, to test the assertion that tuning out would be beneficial for insight and creativity, but zoning out would not. Instead, I found that neither tuning out nor zoning out predicted post-incubation performance on insight or creativity tasks. Post-incubation coins solvers did not tune out more (M = .20, SD = 0.18) than those who never solved (M = .27, SD = 0.21), $t(109)$ = .84, p = .40, 95% CI [-0.09, 0.23], d = 0.35. And in terms of zoning out, post-incubation coins solvers did not differ (*M* = .15, *SD* = 0.10) from those who never solved (*M* = .14, *SD* = 0.15), *t*(109) = -.16, *p* = .88, 95% CI [-0.12, 0.11], *d* = -0.07. Likewise, post-incubation pigpen solvers did not tune out more (*M* = .18, *SD* = 0.20) than those who never solved (*M* = .26, *SD* = 0.21), *t*(108) = 1.31, *p* = .19, 95% CI [-0.04, 0.20], *d* = 0.38. And again, post-incubation pigpen solvers did not differ in zoning out (*M* = .12, *SD* = 0.18) from those who never solved (*M* = .16, *SD* = 0.16), *t*(108) = 0.93, *p* = .36, 95% [-0.05, 0.14], *d* = 0.26. In addition, regardless of what type of scoring was used, post-incubation divergent thinking was not related to tune outs (average *r*(114) = .01, *p* = .95; top-two *r*(114) = -.03, *p* = .75; max-two *r*(114) = .00, *p* = .97) or zone outs (average *r*(114) = -.11, *p* = .24; top-two *r*(114) = -.10, *p* = .27; max-two *r*(114) = .02, *p* = .86).

Pre-incubation solvers for the coins and pigpen problems were excluded in the primary analyses, but will be reported here for completeness. Subjects who solved the coins task before the incubation period did not differ in WMC z-scores (*M* = -.17, *SD* = 0.77) from those who never solved (*M* z-score = .02, *SD* = 0.77), *t*(111) = .72, *p* = .47, 95% CI [-0.34, 0.72], *d* = 0.25, or those who solved after the incubation period (*M* z-score = -.06, *SD* = 1.09), *t*(14) = .25, *p* = .81, 95% CI [-0.88, 1.11], *d* = 0.12. Likewise, in the pigpen problem, subjects who solved the task before the incubation period did not differ in WMC z-scores (*M* = .27, *SD* = 0.65) from those who never solved (*M* z-score = -.11, *SD* = 0.76), *t*(103) = -1.45, *p* = .15, 95% CI [-0.90, 0.14], *d* = -0.53, or those who solved post-incubation (*M* z-score = .60, *SD* = 0.77), *t*(21) = 1.07, *p* = .30, 95% CI [-0.31, 0.98], *d* = 0.47.

Regarding TUT rates, subjects who solved the coins task pre-incubation (*M* TUT rate = .44, *SD* = 0.32) did not differ from those who never solved (*M* TUT rate = .41, *SD* = 0.26), *t*(111) = -.33, *p* = .74, 95% CI [-0.21, 0.15], *d* = -0.10, or those who solved post-incubation (*M* TUT rate = .35, *SD* = 0.16), *t*(12.28) = -.73, *p* = .48, 95% CI [-0.36, 0.18], *d* = -0.35; note that for the latter analysis, the Levene's Test for Equality of Variances was significant, so I used a t-test that does not assume equal variances across groups. Likewise, subjects who solved the pigpen task pre-incubation did not differ in TUT rates (*M* TUT rate = .32, *SD* = 0.25) from those who never solved (*M* TUT rate = .42, *SD* = 0.24), *t*(103) = 1.22, *p* = .23, 95% CI [-0.07, 0.27], *d* = 0.42, or those who solved post-incubation (*M* TUT rate = .30, *SD* = 0.29), *t*(21) = -.16, *p* = .88, 95% CI [-0.26, 0.23], *d* = -0.07.

Pre-incubation creativity scores, which were also omitted in the primary analyses, did not correlate with WMC z-scores (average scoring: *r*(117) = .14, *p* = .14; top-two scoring: *r*(117) = .13, *p* = .18; max-two scoring: *r*(117) = .04, *p* = .66) or with TUT rates during the creativity incubation tasks (average scoring: *r*(117) = -.04, *p* = .65; top-two scoring: *r*(117) = -.10, *p* = .29; max-two scoring: *r*(117) = -.02, *p* = .85).

Although the two analytical problems—crime and Ravens—were used primarily as distractors in this experiment, I report their analyses for completeness. During the crime problem, 17 people (14.4%) solved the problem while 101 people (85.6%) did not. Likewise, 26 people (22.0%) solved the Ravens problem while 92 (78%) did not. Subjects who solved the Ravens problem had a significantly higher WMC z-score (*M* = .31, *SD* = 0.69) than those who did not (*M* = -.08, *SD* = 0.77), *t*(116) = -2.35, *p* = .02, 95% CI [-0.73, -0.06], *d* = -0.54. For the crime problem, however, there was no difference in WMC z-scores between solvers (*M* = .17, *SD* = 0.89) and non-solvers (*M* = -.01, *SD* = 0.76), *t*(116) = -0.89, *p* = .38, 95% CI [-0.58, 0.22], *d* = -0.22.

Discussion

The main hypotheses that Experiment 1 was designed to test, from an individual differences perspective, were as follows: (1) is executive control helpful for insight and creativity? (2) is lack of executive control helpful for insight and creativity? (3) does the correlation between WMC and TUTs vary by context? (4) does concentration moderate the relationship between WMC and TUTs?

In Experiment 1, I found that WMC did not predict post-incubation creativity (as indicated by divergent thinking performance) or post-incubation coins solving, but people who solved the pigpen problem after incubation did have significantly greater WMC than those who never solved. Also, from an individual-differences perspective, mind wandering during an incubation period was not beneficial to either post-incubation insight problem solving or creativity. In addition, the correlation between WMC and TUT rates did not change across tasks, regardless of whether or not subjects were given an experimentally-induced reason to mind wander. Finally, concentration did not moderate the relationship between WMC and TUTs.

The inconsistency in how WMC predicted post-incubation solving in the coins vs. pigpen problem could have been due to the drastic difference in solving rates between the two insight problems. In the pigpen problem, there were 14 post-incubation solvers, but in the coins problem there were only 7. With the rate of post-incubation solving near floor, the coins problem may not have had enough "successful" data points, resulting in an under-powered analysis. Therefore, in Experiment 2, I presented the coins problem using actual coins with the intention that this would increase solving rates. Instead of asking the subjects to draw out the solution, I presented them with large plastic coins arranged in the initial pattern (Figure 1A), and asked them to move two of the coins to achieve the correct answer. Based on work by Ormerod, MacGregor, and Chronicle (2002), who gave their subjects wooden coins to manipulate, I suspected that it would be less of a "leap" to reach the correct solution with three dimensional coins vs. drawing them in a two dimensional format, increasing the number of solvers.

The finding that WMC did not predict post-incubation creativity was somewhat surprising—WMC consistently predicts success in many higher-order cognitive tasks, and is closely related to Gf, which moderately predicts success in creativity tasks. It is possible, however, that because WMC and Gf are not one in the same, the variance in Gf that is not shared with WMC could be responsible for the correlation between creativity and Gf. Alternatively, other research has suggested that WMC may in fact be harmful to creativity, and so it will be important to replicate my findings in Experiment 2 before drawing any conclusions.

From an individual-differences approach, mind wandering during an incubation period was not beneficial to either insight problem solving or creativity (and was actually trending towards harmful in the pigpen task), which contradicts my hypothesis and indirectly conflicts with some past research. It is possible, however, that this null effect could have been due to the type of incubation task that I used. Initially, I had chosen the n-back task, not only because it has been used successfully as an incubation task before (Dijksterhuis & Meurs, 2006), but also because people have reported a significant amount of mind wandering during this task (~50% TUT rates; McVay, Meier, Touron, & Kane, 2013), and so I expected that it would similarly allow for mind wandering in the incubation task. Although the subjects here, once again, reported that they were mind wandering during a substantial number of thought probes (~40%), a meta-analysis of the incubation literature, suggests that an incubation task that requires a low cognitive load is the most beneficial for performance in insight tasks (Sio & Ormerod, 2009). Specifically, previous work that has successfully found an incubation effect has often used science fiction readings as an incubation task (e.g., Smith & Blankenship, 1991). So, in order to give mind wandering its best opportunity to show a benefit to creativity, I based my incubation task for Experiment 2 on several science fiction readings.

I also tested the separate effects of tuning out and zoning out on insight and creativity. Contrary to previous work, which has suggested that tuning out would be beneficial for insight and creativity whereas zoning out would not, I found that *neither* tuning out nor zoning out predicted post-incubation performance on insight or creativity tasks. It will be important to replicate these results before drawing any further conclusions, however, at this point it seems as though mind wandering is not beneficial to insight and creativity, regardless of whether people are aware (e.g., tuning out) or not (e.g., zoning out).

Based on the idea that mind wandering during an incubation period may provide a benefit for creative problem solving, the relationship between WMC and TUTs should vary by context. That is, when there was no experimentally-induced reason to mind wander (i.e., during the stand-alone SART), TUTs should be lower for people with high WMC than for people with low WMC (resulting in a negative correlation), but when there was an experimentally-induced reason to mind wander (i.e., during the n-back incubation task) TUTs might be greater for people with high WMC than for people with low WMC (resulting in a positive correlation). Instead, I found that there was no significant difference between the correlations in the two different contexts and was actually trending in the opposite direction of what I had predicted. This could be because the stand-alone SART was the first task in the first session, when the association between WMC and TUTs is potentially weaker (McVay & Kane, 2009), but it could also be a result of using two different tasks for the two different contexts (i.e., the stand-alone task was always the SART and the incubation task was always the n-back). For this reason, and to confirm that this is not simply a power issue, it will be important to replicate this trend in Experiment 2.

Previous research into daily life has suggested that, without taking concentration into consideration, mind wandering research in the lab may be underestimating the relationship between WMC and TUTs (Kane et al., 2007). Although I did find that concentration predicted mind wandering, it did not moderate the relationship between WMC and mind wandering in Experiment 1. Once again, before drawing any conclusions, it will be important to replicate this finding in the second experiment.

CHAPTER III

EXPERIMENT 2

The procedure was identical to that used in Experiment 1 with two notable exceptions. First, the SART and n-back tasks were replaced with word-by-word prose-reading tasks to measure stand-alone and incubation mind-wandering rates, described below. Second, two personality measures—the openness subscale of the NEO-PI-3 (McCrae & Costa, 2010) and the need for cognition questionnaire (Cacioppo, Petty, & Kao, 1984)—were added to the end of the second session (replacing the Ravens problem). Because the intention behind having the analytical problems included in the sessions was to prevent subjects from noticing a pattern of how to solve the insight problems, it seemed safe to eliminate the Raven's problem since it was the last task in the last session, and would not affect the previously presented insight problems. The openness subscale of the NEO questionnaire was used to measure openness to experience. Most relevant to present purposes, people who score high on openness to experience tend to be intellectually curious and have an increased motivation to be creative (McCrae, 1987). Likewise, I used the need for cognition questionnaire to estimate how often people choose to engage in critical thinking and how much they enjoy it when they do (Cacioppo & Petty, 1982). Therefore, I used these two measures as indicators of intrinsic motivation and valuing the goal of the task—two factors that could positively impact performance on these tasks. Using the openness to experience score and the need for cognition score as moderators in separate analyses, I tested whether the impact of WMC on creativity, insight, or both, would be stronger for subjects who were also high on openness to experience or in need for cognition.

Methods

Subjects

In order to keep my N similar across experiments, while accounting for subjects who would need to be dropped, I aimed to collect data from about 130 people in both sessions of Experiment 2. Once again, undergraduates enrolled in introductory psychology courses at UNCG earned partial credit toward a course requirement for their participation in each of two two-hour sessions. One hundred and fifty-four undergraduates, who had not participated in Experiment 1, completed the first session of Experiment 2; of those, 131 students returned to complete the second session as well. In order to minimize attrition, we gave the majority of the credits upon the completion of the second session.

Reading Tasks

For each incubation task and the stand-alone task for TUT assessment, subjects read a different science fiction story, presented one word at a time on-screen, and responded by pressing the space bar whenever they detected an anomaly: On 5% of the word trials, two adjacent words were swapped (e.g., Bill played fetch *his with* dog), representing an anomaly target. Subjects were asked to respond only after they had seen the second swapped word (e.g., after *with* in my example). Nearly identical to the procedure for the n-back tasks in Experiment 1, each word was presented for 400 ms, followed by a blank screen for 600 ms. This rate was based on Smallwood, McSpadden, and Schooler's (2008) reported average reading time of 304 ms per word in a word-by-word reading comprehension study. Presenting words at a constant pace helped to control for differences in self-paced reading speed and ensured that all subjects had the same incubation time. During this 20 min task, subjects saw approximately 1200 words, 61 of which were targets; thought probes appeared immediately after approximately 60% of the targets. These 35 thought probes are identical in format to, and comparable in number to, the probes used in the Experiment 1 SART (60 probes) and n-back tasks (21 probes). After each story, subjects answered five multiple-choice questions about the story, via mouse click, to motivate their reading for comprehension.

Non-Cognitive Assessments

We used two personality inventories—openness and need for cognition—to measure intellectual and creative motivation, but I combined them into one computer-presented questionnaire. Questions from the need for cognition measure were dispersed evenly throughout the openness questions, using a repeating pattern of two openness items, followed by one need for cognition item, followed by three openness items, followed by one need for cognition item, followed by three openness items, followed by one more need for cognition item before looping back to the beginning of the sequence with two openness items, etc. until all 66 items were included. Both questionnaires asked subjects to respond using a five-option Likert-type scale, ranging from "Strongly Agree" to "Strongly Disagree" (i.e., Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree).

Openness. This questionnaire, which was taken from the NEO-PI-3 (McCrae & Costa, 2010), is made up of six facets—fantasy (e.g., *I enjoy concentrating on a fantasy or daydream and exploring all its possibilities, letting it grow and develop*), aesthetics (e.g., *I am intrigued by the patterns I find in art and nature*), feelings (e.g., *I experience a wide range of emotions or feelings*), actions (e.g., *I think it's interesting to learn and develop new hobbies*), ideas (e.g., *I often enjoy playing with theories or abstract ideas*), and values (e.g., *I consider myself broad-minded and tolerant of other people's lifestyles*)—with eight questions defining each and half reverse-coded. Although I was primarily interested in the general openness construct, of secondary interest to me was the fantasy facet, which, at face value, appears to measure propensity to mind wander.

Need for cognition. I used the short form of the need for cognition questionnaire (Cacioppo, Petty, & Kao; 1984), which consists of 18 questions (nine of which are reverse-coded). Although the original questionnaire used a -4 to +4 Likert-type scale, I used a -2 to +2 scale in order to keep the responses consistent with the openness measure in the combined questionnaire (i.e., Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree).

Results

Subjects

Data from subjects were omitted for processing-portion errors on either span task (14 people) and for one person who was older than my target, young adult, age group of 18-30. Using the same criterion as in Experiment 1, outliers were determined by collapsing non-target accuracy across all reading tasks (both stand-alone and incubation), but here, nobody had an accuracy of 75% or less and so data from all remaining subjects were retained (mean accuracy and reaction times can be found in Appendix C). A total of 116 people were included for analyses, ranging in age from 18-28 years (*M* = 18.88, 66.4% female). By self-report, the final sample's racial composition was 56.0% White, 31.0% Black, 3.4% Asian, 5.2% Multiracial, 1.7% Native American or Alaskan Native, and 2.6% Other. Additionally, regarding self-reported ethnicity, 4.3% self-identified as Latino/Hispanic.

Primary Analyses

Descriptive statistics for WMC and TUT measures can be found in Table 7, and descriptive statistics for the non-cognitive measures can be found in Table 8. The respective correlation matrices for these measures can be found in Tables 9 and 10. As in Experiment 1, I collapsed across the two WMC tasks (operation span and symmetry span) to create one WMC z-score composite measure. Average WMC z-scores were lower in Experiment 2 (*M* = -0.2076, *SD* = 0.84), compared to Experiment 1 (*M* = -0.0005, *SD* = 0.78), $t(234) = 1.960$, $p = .051$, 95% CI [0.00, 0.42], $d = 0.25$. Across all reading tasks, subjects were mind wandering approximately 30% of the time, regardless of whether the task was stand-alone or as an incubation period. Just as in Experiment 1, a one-way analysis of variance (ANOVA) indicated no statistical differences in average TUT rates across any of the five different tasks, $F(4, 573) = 1.165$, $p = .325$, $MSE = 0.058$, $\eta_p^2 = .008$). Furthermore, TUT rates correlated significantly across all across all pairwise tasks, *r*s = .40 to .72 (see Table 9), indicating that mind wandering is also consistent across people. Overall, TUT rates were significantly lower than in Experiment

1, and this remained true whether I took the average TUT rate across all five tasks, $t(230) = 4.582$, $p < .001$, 95% CI [0.06, 0.16], $d = 0.60$, or compared the tasks one-by-one (stand-alone: $t(233) = 4.175$, $p < .001$, 95% CI [0.06, 0.16], $d = 0.54$; coins: $t(233) = 4.647$, $p < .001$, 95% CI [0.09, 0.21], $d = 0.61$; pigpen: $t(232) = 3.677$, $p < .001$, 95% CI [0.06, 0.19], $d = 0.48$; brick: $t(234) = 3.220$, $p = .001$, 95% CI [0.04, 0.16], $d = 0.42$; or knife: $t(233) = 2.861$, $p = .005$, 95% CI [0.03, 0.16], $d = 0.37$).

Table 7

Experiment 2 Descriptive Statistics for Working Memory Capacity Tasks and Mind Wandering

	N	Mean	SD	Min	Max	Skew	Kurtosis
Sspan Z	116	-0.325	1.040	-2.615	1.959	-0.245	-0.682
Ospan Z	116	-0.091	1.022	-3.145	1.546	-0.832	0.452
WMC Z	116	-0.208	0.841	-2.783	1.585	-0.519	0.069
Stand-alone TO	116	0.127	0.117	0	0.571	1.081	1.065
Stand-alone ZO	116	0.128	0.146	0	0.722	1.462	2.197
Stand-alone TUT	116	0.256	0.212	0	0.861	0.951	0.239
Coins TO	115	0.139	0.146	0	0.657	1.451	2.234
Coins ZO	115	0.126	0.162	0	0.943	2.050	5.612
Coins TUT	115	0.265	0.227	0	1	0.962	0.481
PigPen TO	115	0.146	0.173	0	0.943	2.135	5.464
PigPen ZO	115	0.132	0.168	0	0.944	2.262	6.698
PigPen TUT	115	0.278	0.251	0	0.944	1.141	0.438
Brick TO	116	0.151	0.183	0	0.914	2.108	5.075
Brick ZO	116	0.145	0.167	0	0.857	1.742	3.607
Brick TUT	116	0.296	0.245	0	1	0.943	0.190
Knife TO	116	0.137	0.148	0	0.829	1.670	3.733
Knife ZO	116	0.178	0.199	0	0.943	1.433	1.912
Knife TUT	116	0.315	0.262	0	1	0.733	-0.250

Note. Z = Z scores based on database of over 3000 people; WMC = working memory capacity; TO = tune outs; ZO = zone outs; TUT = task-unrelated thoughts.

Table 8

Experiment 2 Descriptive Statistics for Non-Cognitive Variables

	N	Mean	SD	Min	Max	Skew	Kurtosis
Fantasy Facet	116	0.610	1.106	-1.543	3.022	-0.037	-0.714
Aesthetics Facet	116	0.747	0.996	-2.000	2.561	-0.300	-0.421
Feelings Facet	116	0.609	1.065	-2.049	2.829	-0.479	0.198
Actions Facet	116	0.485	1.179	-2.194	3.917	0.386	0.117
Ideas Facet	116	0.438	0.847	-1.259	2.630	0.274	0.078
Values Facet	116	0.358	1.077	-2.643	2.833	-0.090	-0.017
Openness	116	0.816	1.047	-2.242	3.134	-0.011	0.121
Need for Cognition	116	0.000	1.000	-2.688	2.852	0.138	-0.065

Note. All measures are z-scores; openness measures are based on normed means and standard deviations provided by McCrae & Costa (2010); need for cognition was calculated internally.

Table 9

Experiment 2 Correlations Among Working Memory Capacity Tasks and Mind Wandering

	1	2	3	4	5	6	7	8	9	10
1. Sspan Z	1.00									
2. Ospan Z	.33**	1.00								
3. WMC Z	.82**	.81**	1.00							
4. Stand-alone TOs	-.03	-.01	-.03	1.00						
5. Stand-alone ZOs	.00	.28**	.17	.30**	1.00					
6. Stand-alone TUTs	-.02	.19*	.10	.76**	.85**	1.00				
7. Coins TOs	-.02	-.05	-.04	.40**	.07	.27**	1.00			
8. Coins ZOs	-.06	-.02	-.05	.33**	.48**	.51**	.09	1.00		
9. Coins TUTs	-.05	-.04	-.06	.49**	.39**	.53**	.70**	.77**	1.00	
10. Pigpen TOs	.00	.10	.06	.35**	.10	.27**	.58**	.17	.50**	1.00
11. Pigpen ZOs	.01	.18	.12	.26**	.60**	.56**	.12	.51**	.44**	.08
12. Pigpen TUTs	.01	.19*	.12	.42**	.48**	.56**	.48**	.46**	.64**	.75**
13. Brick TOs	.06	.05	.07	.33**	.06	.22*	.71**	.09	.52**	.62**
14. Brick ZOs	-.02	.06	.03	.32**	.59**	.58**	-.01	.67**	.47**	.09
15. Brick TUTs	.04	.08	.07	.47**	.44**	.56**	.52**	.52**	.71**	.53**
16. Knife TOs	.01	.04	.03	.42**	.11	.31**	.43**	.21*	.43**	.65**
17. Knife ZOs	.02	.15	.10	.27**	.60**	.56**	.04	.41**	.32**	.09
18. Knife TUTs	.02	.13	.09	.44**	.52**	.60**	.27**	.43**	.48**	.44**
19. Incubation TOs	.01	.05	.04	.46**	.10	.32**	.83**	.16	.65**	.88**
20. Incubation ZOs	-.01	.13	.07	.38**	.72**	.70**	.06	.81**	.62**	.14
21. Incubation TUTs	.00	.12	.07	.55**	.54**	.67**	.58**	.64**	.84**	.67**

Table 9 (continued)

	11	12	13	14	15	16	17	18	19	20
1. Sspan Z										
2. Ospan Z										
3. WMC Z										
4. Stand-alone TOs										
5. Stand-alone ZOs										
6. Stand-alone TUTs										
7. Coins TOs										
8. Coins ZOs										
9. Coins TUTs										
10. Pigpen TOs										
11. Pigpen ZOs	1.00									
12. Pigpen TUTs	.73**	1.00								
13. Brick TOs	.14	.52**	1.00							
14. Brick ZOs	.33**	.29**	-.02	1.00						
15. Brick TUTs	.33**	.59**	.73**	.67**	1.00					
16. Knife TOs	.20*	.58**	.36**	.17	.39**	1.00				
17. Knife ZOs	.68**	.52**	.05	.30**	.24**	.13	1.00			
18. Knife TUTs	.63**	.72**	.24**	.33**	.40**	.66**	.83**	1.00		
19. Incubation TOs	.16	.71**	.84**	.06	.67**	.74**	.09	.48**	1.00	
20. Incubation ZOs	.81**	.64**	.08	.72**	.54**	.23*	.79**	.73**	.15	1.00
21. Incubation TUTs	.65**	.89**	.60**	.52**	.80**	.63**	.58**	.80**	.75**	.76**

Note. Z = Z scores based on database of over 3000 people; WMC = working memory capacity; TO = tune outs; ZO = zone outs; TUT = task-unrelated thoughts

* $p < .05$; ** $p < .01$

Table 10

Experiment 2 Correlations of the Non-Cognitive Variables with the Working Memory Capacity Tasks and Mind-Wandering Rates

	Fantasy Facet	Aesthetics Facet	Feelings Facet	Actions Facet	Ideas Facet	Values Facet	Openness	Need for Cognition
1. Sspan Z	.06	.07	-.14	-.24**	.02	-.05	-.06	-.17
2. Ospan Z	-.05	-.14	-.14	-.25**	-.18	-.15	-.22*	-.16
3. WMC Z	.01	-.05	-.17	-.30**	-.10	-.12	-.16	-.20*
4. Stand-alone TUTs	.14	-.03	.03	-.04	-.04	.04	.02	-.01
5. Coins TUTs	.13	.10	.07	.19*	-.03	.01	.12	-.01
6. Pigpen TUTs	.03	.07	-.05	.12	-.12	.03	.02	-.09
7. Brick TUTs	-.01	.01	.05	-.05	-.20*	-.08	-.07	-.12
8. Knife TUTs	.15	.13	-.01	.15	-.10	.17	.13	-.13
9. Incubation TUTs	.08	.08	.00	.12	-.13	.04	.05	-.12
10. Average Pre Inc DT	.28**	.22*	.18	.28**	.06	.20*	.30**	.02
11. Average Post Inc DT	.25*	.13	.17	.21*	.08	.21*	.26**	.00
12. Top-Two Pre Inc DT	.18	.10	.15	.14	-.04	.05	.14	-.03
13. Top-Two Post Inc DT	.18	.01	.17	.25**	.08	.14	.19*	.00
14. Max-Two Pre Inc DT	.24*	.16	.17	.20*	-.01	.03	.20*	.02
15. Max-Two Post Inc DT	.20*	-.03	.09	.15	.04	.06	.12	.00

Note. Z = Z scores based on database of over 3000 people; WMC = working memory capacity; TO = tune outs; ZO = zone outs; TUT = task-unrelated thoughts; DT = divergent thinking.
* $p < .05$; ** $p < .01$

Changing the coins insight problem to a three-dimensional, manipulable format increased the number of solvers, as expected, yielding 15 pre-incubation solvers (12.9%), 22 post-incubation solvers (19.0%), and 79 people who never solved the coins problem (68.1%). Unfortunately, in the pigpen problem, rates of solution dropped substantially from Experiment 1, despite its identical presentation: There were 9 pre-incubation solvers (7.8%), 7 post-incubation solvers (6.0%), and 99 who never solved (85.3%). In addition, there was one person who had seen the pigpen problem before, and was therefore excluded for any pigpen analyses (0.9%). Once again, people who solved either insight problem pre-incubation were excluded from the relevant, primary analyses.

As stated before, inter-rater reliability across the two experiments was quite good for both brick and knife (see Table 1). Also, overall creativity (collapsed across pre- and post-incubation) in the brick task was correlated with overall creativity in the knife task, regardless of using average, $r(111) = .55$, $p < .001$, top-two, $r(112) = .40$, $p < .001$, or max-two scoring, $r(112) = .50$, $p < .001$. As in Experiment 1, brick and knife scores were averaged together to create one creativity measure.

Is executive control helpful for insight and creativity? Similar to the results from Experiment 1, subjects who solved the coins problem after incubation did not have a significantly different WMC z-score than did those who never solved the coins problem (M = -.28 [SD = 0.89] & -.25 [SD = 0.85], respectively), $t(99) = 0.15$, $p = .88$, 95% CI [-0.38, 0.44], $d = 0.04$. For the pigpen problem, I replicated the Experiment 1 pattern, finding that WMC z-scores were higher for post-incubation pigpen solvers (M = .23, SD = 0.57) than for subjects who never solved (M = -.25, SD = 0.87), but the difference was not significant here, *t*(104) = -1.44, *p* = .15, 95% CI [-1.15, 0.18], *d* = -0.65. In addition, WMC was, again, not a significant predictor of post-incubation creativity (average *r*(107) = .04, *p* = .65; top-two *r*(107) = -.02, *p* = .81; max-two *r*(107) = .01, *p* = .94; see Figure 5).

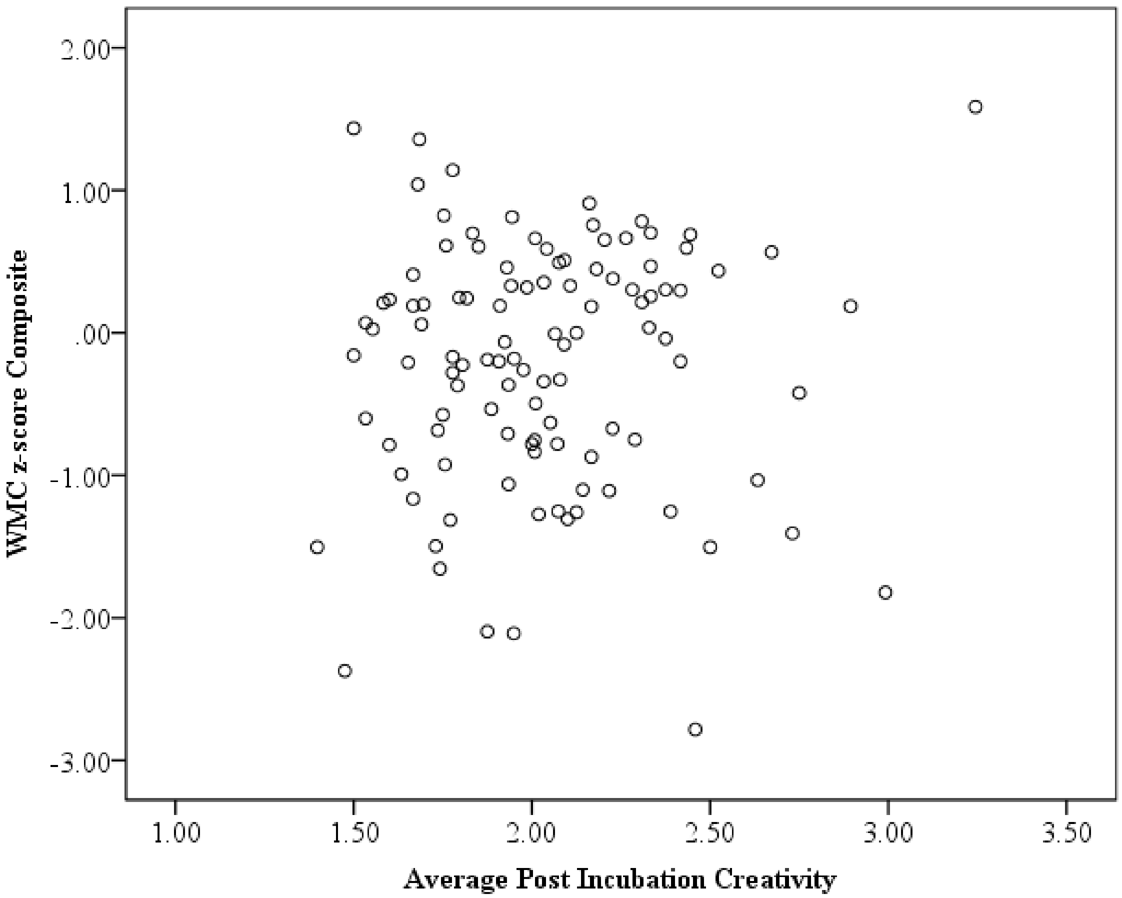

Figure 5. Experiment 2 Correlation Scatter between WMC and Creativity (Average Scoring).

Is lack of executive control helpful for insight and creativity? Similar to the results from Experiment 1, subjects who solved the coins problem after incubation did not mind wander more during the coins incubation task than those who never solved the coins problem (*M* TUT rates = .25 [*SD* = 0.28] and .28 [*SD* = 0.22], respectively), $t(99) = 0.55$, $p = .59$, 95% CI [-0.08, 0.14], $d = 0.12$. And, if anything, subjects who solved the pigpen problem after incubation mind wandered *less* (*M* TUT rate = .14, *SD* = 0.13) than those who never solved (*M* TUT rate = .29, *SD* = 0.26), but not significantly so, $t(104) = 1.49$, $p = .14$, 95% CI [-0.05, 0.35], $d = 0.72$.

Just as I found in Experiment 1, mind wandering during the incubation task was not beneficial for post-incubation creativity scores, regardless of using an average, $r(107) = .04$, $p = .67$, top-two, $r(107) = -.07$, $p = .47$, or max-two scoring method, $r(107) = .02$, $p = .89$, (see Figure 6).

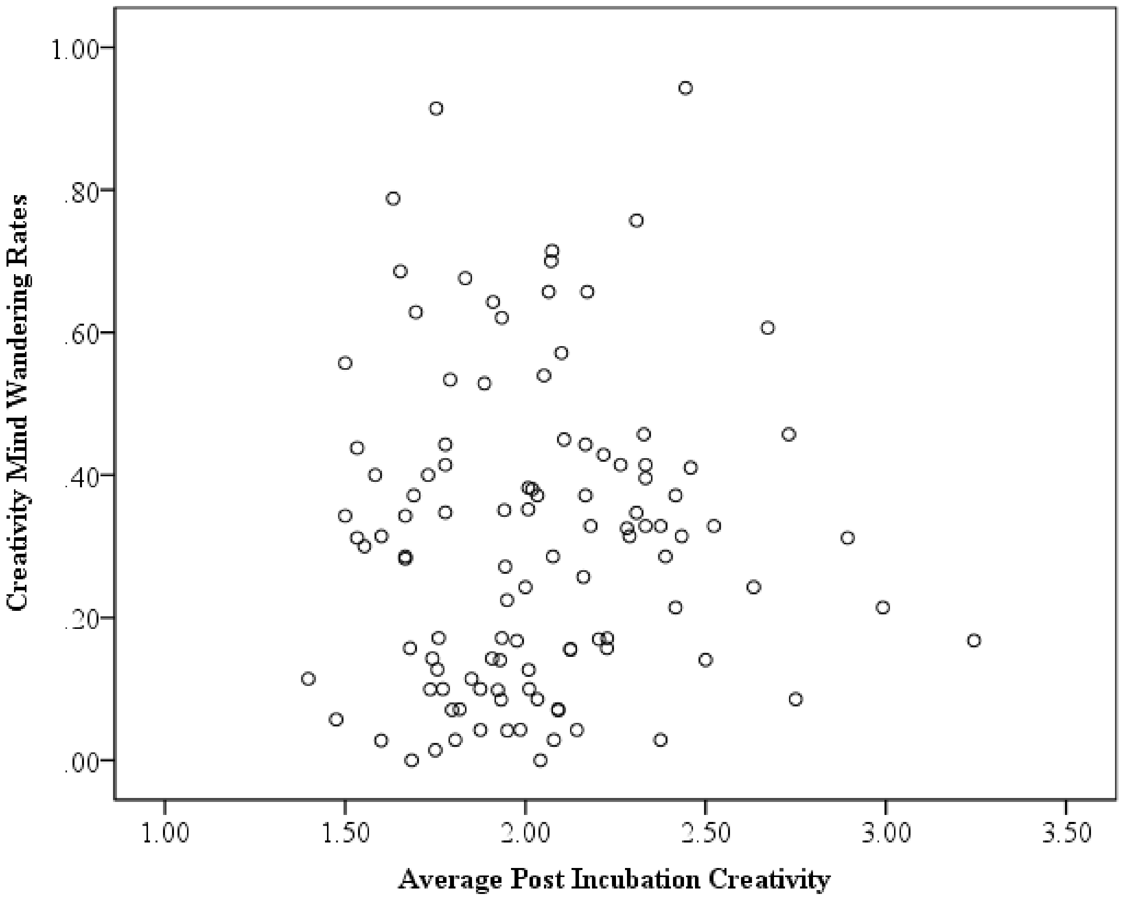

Figure 6. Experiment 2 Correlation Scatter between TUT Rates and Creativity (Average Scoring).

Once again, I tested whether the interaction between WMC and TUT rates could predict insight or creativity, with the possibility that people high in WMC *and* TUT rates may be more insightful or creative. Logistic regressions testing whether this interaction predicted insight can be found in Table 11. Unlike Experiment 1 (where post-incubation pigpen solvers were high in WMC and low in TUT rates), I found no significant interaction for either coins or pigpen problems. For the creativity outcome, a regression also indicated no significant WMC x TUT rate interaction (see Table 12).

Table 11

Experiment 2 WMC x TUT Logistic Regression on Insight Problem Solving

	B	*SE*	*Wald*	*p*	*Exp(B)*	*95% CI for Exp(B)* Lower	Upper
Coins Predictors							
Block 1							
WMC z-score	-0.045	0.239	0.036	.850	0.956	0.598	1.527
Coins TUT Rate	-0.139	0.249	0.312	.576	0.870	0.534	1.418
Constant	-1.281	0.242	27.928	.000	0.278		
Block 2							
WMC x TUT	0.174	0.229	0.577	.447	1.190	0.760	1.865
Pigpen Predictors							
Block 1							
WMC z-score	0.725	0.478	2.301	.129	2.064	0.809	5.265
Pigpen TUT Rate	-0.949	0.639	2.202	.138	0.387	0.111	1.356
Constant	-3.098	0.569	29.673	.000	0.045		
Block 2							
WMC x TUT	0.036	0.812	0.002	.964	1.037	0.211	5.090

Note. WMC = working memory capacity; TUT = task-unrelated thoughts; Coins N = 101, Pigpen N = 106; all predictors are z-scores (for centering purposes).

Table 12

Experiment 2 WMC x TUT Hierarchical Linear Regression on Creativity

Predictors	*B*	*SE*	*β*	*t*	*p*	*95% CI for B* Lower	Upper
Block 1 ($\Delta R^2 = .003$)							
WMC z-score	0.014	0.033	0.040	0.413	.680	-0.052	0.079
Creativity TUT Rate	0.012	0.033	0.037	0.380	.705	-0.053	0.078
Constant	2.032	0.033		61.663	.000	1.967	2.097
Block 2 ($\Delta R^2 = .007$)							
WMC x TUT	-0.029	0.034	-0.084	-0.853	.396	-0.097	0.039

Note. WMC = working memory capacity; TUT = task-unrelated thoughts; N = 109; all predictors are z-scores (for centering purposes).

Cross-experimental analyses of pigpen solving and divergent thinking. Because my results regarding the pigpen problem in Experiment 2 were somewhat inconsistent (in terms of statistical significance) with those from Experiment 1, I collapsed the data sets across my two experiments to increase power, and re-ran the analyses. Here, post-incubation solvers' WMC scores (M = 0.48, SD = 0.72, N = 21) were significantly higher, $t(214) = -3.54$, $p < .001$, 95% CI [-1.03, -0.29], d = -0.86, than were non-solvers' (M = -0.18, SD = 0.82, N = 195). Also, post-incubation pigpen solvers tended to mind wander less (M TUT rate = .25, SD = 0.25) than did people who never solved (M TUT rate = .36, SD = 0.26), but not significantly so, $t(214) = 1.82$, p = .07, 95% CI [-0.01, 0.23], d = 0.42. Finally, I ran a logistic regression to test whether the interaction between WMC and TUT rates predicted solving in the pigpen insight problem. WMC and pigpen TUT rates were entered into the regression at Step 1 and the interaction between the two (WMC x TUT) was entered at Step 2. The interaction did predict pigpen solving in the same direction as in Experiment 1, but not significantly (p = .08; see Table 13).

Table 13

Cross-Experimental WMC x TUT Logistic Regression on Pigpen Insight Problem Solving

						95% CI for Exp(B)	
	B	*SE*	*Wald*	*p*	*Exp(B)*	*Lower*	*Upper*
<u>Pigpen Predictors</u>							
Block 1							
WMC z-score	0.986	0.301	10.740	.001	2.681	1.486	4.834
Pigpen TUT Rate	-0.483	0.277	3.034	.082	0.617	0.359	1.062
Constant	-2.637	0.318	68.800	.000	0.072		
Block 2							
WMC x TUT	-0.538	0.310	3.007	.083	0.584	0.318	1.073

Note. WMC = working memory capacity; TUT = task-unrelated thought; N = 216; Pigpen TUT Rates were centered and WMC z-scores were re-centered after combining Experiment 1 and 2.

We were also able to run cross-experimental analyses on creativity because the tasks did not change across experiments. To no surprise, when collapsing the data from the divergent thinking tasks across experiments, I replicated the results found in Experiment 1 and 2. Regardless of whether average, top-two, or max-two scoring was used, neither WMC ($r(223) = .03$, $p = .64$; $r(223) = .03$, $p = .71$; $r(223) = .02$, $p = .72$, respectively) nor TUT rates during divergent-thinking incubation ($r(223) = -.01$, $p = .90$; $r(223) = -.08$, $p = .25$; $r(223) = .02$, $p = .77$, respectively) predicted post-incubation creativity. In addition, using a hierarchical multiple regression with WMC and creativity TUT rates entered at Step 1 and an interaction of WMC x TUT entered at Step 2, nothing significantly predicted post-incubation creativity (see Table 14).

Table 14

Cross-Experimental WMC x TUT Hierarchical Linear Regression on Creativity

						95% CI for B	
Predictors	*B*	*SE*	*β*	*t*	*p*	*Lower*	*Upper*
Block 1 (Δ R^2 = .001)							
WMC z-score	0.010	0.022	0.032	0.471	.638	-0.032	0.053
Creativity TUT Rate	-0.003	0.022	-0.010	-0.144	.885	-0.046	0.039
Constant	2.038	0.022		94.246	.000	1.995	2.081
Block 2 (Δ R^2 = .000)							
WMC x TUT	-0.004	0.022	-0.013	-0.187	.852	-0.048	0.040

Note. WMC = working memory capacity; TUT = task-unrelated thought; N = 225; Creativity TUT Rates were centered and WMC z-scores were re-centered after combining Experiment 1 and 2.

Does the correlation between WMC and TUTs vary by context? Once again, I tested whether or not the correlation between WMC and TUT rate varied as a function of task type (i.e., stand-alone vs. incubation). Comparing the correlation between WMC and TUT rate during the stand-alone task (r = .10) to the correlation between WMC and TUT rate during the incubation tasks (r = .07), I found that there was no statistically significant difference in magnitude between the two, $t(112) = 0.41$, $p = .68$.

Does concentration moderate the relationship between WMC and TUTs? Using the same multilevel model as in Experiment 1, and collapsing across all of the reading tasks, I included a fixed effect for concentration at Level 1 and fixed effects for WMC on the intercept and on the concentration slope at Level 2 to predict TUTs (0 = on task, 1 = off task) on any given trial. Also like Experiment 1, I added a random effect for the intercept. Once again, I found that concentration negatively predicted TUTs ($\beta_{10} = -1.10$, $t(17383) = -17.67$, $p < .001$; see Table 15), but WMC did not significantly predict TUTs

(β_{01} = -0.02, t(114) = -0.09, p = .93), and concentration did not moderate the association between WMC and TUTs (β_{11} = -0.11, t(17383) = -1.22, p = .23; see Figure 7). Finally, a significant person-level random effect on the intercept suggests that there remain individual differences in TUT rates beyond what WMC and concentration can explain.

Table 15

Experiment 2 WMC x Concentration Hierarchical Linear Model on Mind Wandering

Fixed Effect	Coefficient	*SE*	*T-ratio*	*d.f.*	*p*
Intercept, β_{00}	-0.988	0.171	-5.767	114	.000
WMC, β_{01}	-0.018	0.204	-0.090	114	.929
Concentration, β_{10}	-1.097	0.062	-17.666	17383	.000
WMC, β_{11}	-0.112	0.092	-1.215	17383	.225

Random Effect	χ^2	*SD*	*Variance Component*	*d.f.*	*p*
Intercept, r_0	4161.386	1.807	3.264	114	.000

Note. WMC = working memory capacity; calculated by averaging across operation span and symmetry span z-scores. In this analysis, the main effects of WMC and concentration are represented by β_{01} and β_{10}, respectively, and the interaction between the two is represented by β_{11}.

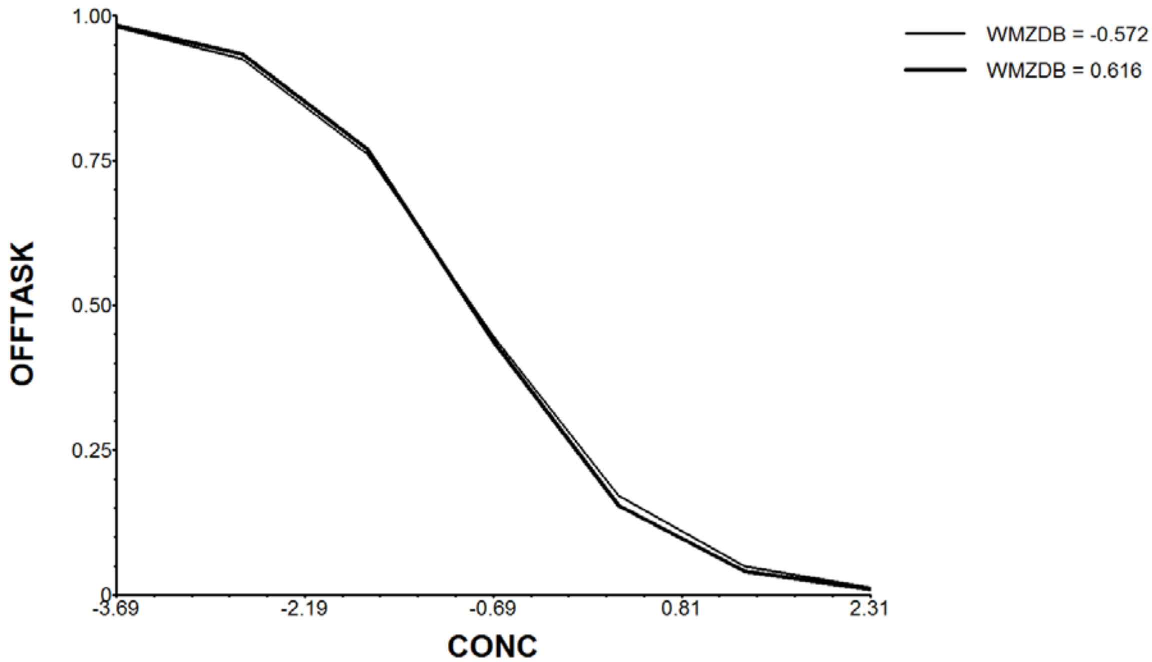

Figure 7. The Relation between Mind Wandering and Concentration for Subjects with High and Low WMC in Experiment 2. The lines represent the means of the within-person slopes for subjects in the top quartile of WMC (thicker line) and the bottom quartile of WMC (thinner line). Mind wandering is on the y-axis, where 0 = on-task and 1 = off-task. Concentration is group-centered on the x-axis, where negative values indicate less concentration and positive values indicate more concentration.

Does intellectual motivation moderate the relationship between WMC and insight or creativity? Before testing my hypothesis, I first looked at how these measures relate to one another. The openness subscale of the NEO was positively correlated with the need for cognition questionnaire, $r(114) = .42$, $p < .001$, but trended negatively with WMC, $r(114) = -.16$, $p = .08$. In addition, need for cognition was significantly and negatively correlated with WMC, $r(114) = -.20$, $p = .03$. Regarding creativity, openness was significantly correlated with the average score of post-incubation creativity, $r(107) = .26$, $p < .01$ (see Figure 8), and the top-two score, $r(107) = .19$, $p < .05$, but not the max-two score, $r(107) = .12$, $p = .22$. The need for cognition questionnaire, on the other hand, did not correlate significantly with any measure of post-incubation creativity (average $r(107) = .00$, $p = 1.00$; top-two $r(107) = .00$, $p = 1.00$; max-two $r(102) = .00$, $p = .99$; see Figure 9). Contrary to predictions, scores on the openness subscale were not predictive of post-incubation solving in the coins task (*M* z-score for solvers = 1.07, *SD* = 0.99; *M* z-score for non-solvers = .81, *SD* = 1.04), $t(99) = -1.07$, $p = .29$, 95% CI [-0.76, 0.23], $d = -0.26$, or the pigpen task (*M* z-score for solvers = .47, *SD* = 1.60; *M* z-score for non-solvers = .82, *SD* = 1.01), $t(6.346) = .57$, $p = .59$, 95% CI [-1.13, 1.83], $d = 0.26$; note that equal variances were not assumed. need for cognition scores were not predictive of post-incubation solving in the pigpen task (*M* z-score for solvers = .49, *SD* = 1.27; *M* z-score for non-solvers = -.01, *SD* = 0.98), $t(104) = -1.30$, $p = .20$, 95% CI [-1.28, 0.27], $d = -0.45$, but they were indicative of post-incubation coins solving (*M* z-score for solvers = .39, *SD* = 0.99; *M* z-score for non-solvers = -.10, *SD* = 0.99), $t(99) = -2.04$, $p = .04$, 95% CI [-0.96, -0.01], $d = -0.49$.

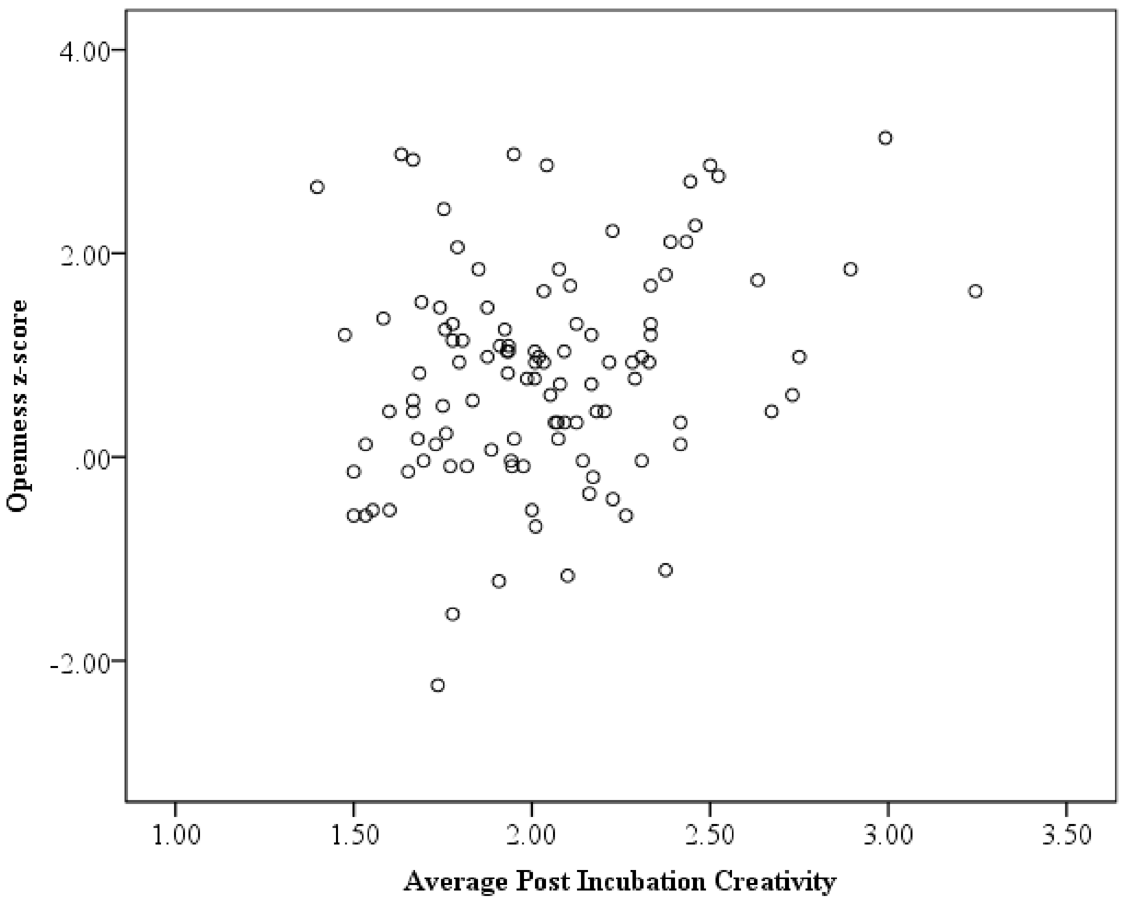

Figure 8. Experiment 2 Correlation Scatter between Openness and Creativity (Average Scoring).

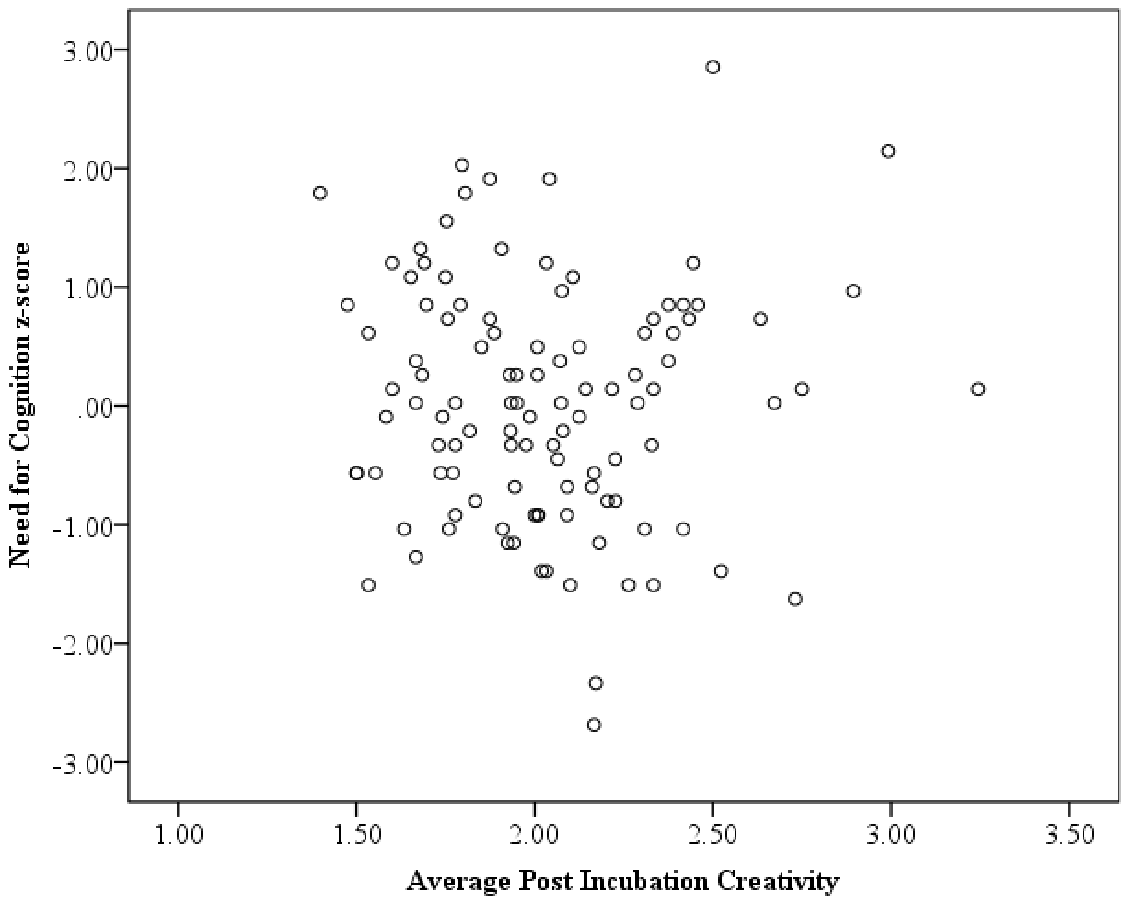

Figure 9. Experiment 2 Correlation Scatter between Need for Cognition and Creativity (Average Scoring).

To test my hypothesis that openness and need for cognition might interact with WMC to predict insight and creativity, I ran several regressions (see Tables 16–19). In my first hierarchical linear regression, I entered WMC z-score and openness in Block 1 to predict average post-incubation creativity, and added in the WMC x openness interaction at Block 2 (Table 16). Likewise, in my second hierarchical linear regression, I entered WMC z-score and need for cognition in Block 1 and added the WMC x need for cognition interaction in Block 2 (Table 17). Contrary to my predictions, neither openness nor need for cognition moderated the relationship between WMC and average post-incubation creativity. In order to test the influence of these same interactions on insight problem solving, I ran separate logistic regressions for coins and pigpen (see Tables 18–19). Similar to the regressions for creativity, WMC z-scores and the personality measure (either openness or need for cognition) were entered at Block 1 and the interaction between the two was entered at Block 2. Once again, the interaction between WMC and openness did not predict post-incubation solving in either coins or pigpen problem solving, and neither did the interaction between WMC and need for cognition.

Table 16

Experiment 2 WMC x Openness Hierarchical Linear Regression on Creativity

						95% CI for B	
Predictors	*B*	*SE*	*β*	*t*	*p*	*Lower*	*Upper*
Block 1 (ΔR^2 = .072)							
WMC z-score	0.029	0.032	0.087	0.916	.362	-0.034	0.093
Openness	0.091	0.032	0.268	2.832	.006	0.027	0.155
Constant	2.032	0.032		63.907	.000	1.969	2.095
Block 2 (ΔR^2 = .000)							
WMC x Openness	0.002	0.031	0.008	0.077	.939	-0.058	0.063

Note. WMC = working memory capacity; N = 109

Table 17

Experiment 2 WMC x Need for Cognition Hierarchical Linear Regression on Creativity

Predictors	*B*	*SE*	*β*	*t*	*p*	*95% CI for B* Lower	Upper
Block 1 (Δ R^2 = .002)							
WMC z-score	0.015	0.033	0.046	0.461	.645	-0.051	0.082
Need for Cognition	0.003	0.034	0.008	0.085	.933	-0.064	0.069
Constant	2.032	0.033		61.617	.000	1.967	2.097
Block 2 (Δ R^2 = .006)							
WMC x Need for Cog	-0.026	0.031	-0.083	-0.826	.411	-0.087	0.036

Note. WMC = working memory capacity; Need for Cog = need for cognition; N = 109.

Table 18

Experiment 2 WMC x Openness Logistic Regression on Insight Problem Solving

	B	*SE*	*Wald*	*p*	*Exp(B)*	*95% CI for Exp(B)* Lower	Upper
Coins Predictors							
Block 1							
WMC z-score	0.000	0.236	0.000	1.000	1.000	0.630	1.588
Openness	0.268	0.253	1.118	.290	1.307	0.795	2.149
Constant	-1.310	0.247	28.075	.000	0.270		
Block 2							
WMC x Openness	0.270	0.249	1.175	.278	1.310	0.804	2.135
Pigpen Predictors							
Block 1							
WMC z-score	0.634	0.466	1.848	.174	1.885	0.756	4.704
Openness	-0.290	0.413	0.492	.483	0.749	0.333	1.681
Constant	-2.862	0.473	36.611	.000	0.057		
Block 2							
WMC x Openness	-0.067	0.433	0.024	.877	0.935	0.401	2.184

Note. WMC = working memory capacity; Coins N = 101, Pigpen N = 106; all predictors are z-scores (for centering purposes).

Table 19

Experiment 2 WMC x Need for Cognition Logistic Regression on Insight Problem Solving

						95% CI for Exp(B)	
	B	*SE*	*Wald*	*p*	*Exp(B)*	*Lower*	*Upper*
<u>Coins Predictors</u>							
Block 1							
WMC z-score	0.075	0.244	0.095	.758	1.078	0.668	1.741
Need for Cog	0.519	0.260	3.973	.046	1.681	1.009	2.800
Constant	-1.348	0.256	27.769	.000	0.260		
Block 2							
WMC x Need for Cog	0.152	0.240	0.403	.526	1.165	0.727	1.865
<u>Pigpen Predictors</u>							
Block 1							
WMC z-score	0.754	0.455	2.744	.098	2.126	0.871	5.188
Need for Cog	0.662	0.422	2.459	.117	1.938	0.848	4.430
Constant	-2.990	0.508	34.597	.000	0.050		
Block 2							
WMC x Need for Cog	0.061	0.448	0.019	.891	1.063	0.442	2.560

Note. WMC = working memory capacity; Need for Cog = need for cognition; Coins N = 101, Pigpen N = 106; all predictors are z-scores (for centering purposes).

Secondary Analyses

Just as I did in Experiment 1, I tested the idea that tuning out would be beneficial for insight and creativity, but zoning out would not. Once again, I found that post-incubation coins solvers did not tune out more (M = .13, SD = 0.18) than those who never solved (M = .15, SD = 0.14), $t(99) = 0.54$, p = .59, 95% CI [-0.05, 0.09], d = 0.12. And in terms of zoning out, post-incubation coins solvers did not differ (M = .12, SD = 0.21) from those who never solved (M = .13, SD = 0.15), $t(99) = 0.27$, p = .79, 95% CI [-0.07, 0.09], d = 0.06.

Unlike Experiment 1, and contrary to Smallwood, McSpadden, and Schooler's (2008) theory, post-incubation pigpen solvers actually tuned out significantly *less* (M = .05, SD = 0.04) than those who never solved (M = .15, SD = .18), $t(33.24) = 4.43$, $p <$ .001, 95% [0.06, 0.15], d = 0.79 (note that equal variances were not assumed). In terms of zoning out, however, post-incubation pigpen solvers did not differ (M = .09, SD = 0.11) from those who never solved (M = .14, SD = 0.18), $t(104) = 0.66$, p = 0.51, 95% [-0.09, 0.18], d = 0.31. Parallel to Experiment 1 results, regardless of what type of scoring was used, post-incubation divergent thinking was not related to tune outs (average $r(107)$ = -.08, p = .42; top-two $r(107)$ = -.07, p = .46; max-two $r(107)$ = -.08, p = .42) or zone outs (average $r(107)$ = .14, p = .16; top-two $r(107)$ = -.04, p = .71; max-two $r(107)$ = .10, p = .31).

Cross-experimental analyses of the pigpen problem revealed that, overall, people who solve after the incubation period did *not* significantly differ in tune outs (M = .14, SD = 0.18) from those who never solved (M = .21, SD = 0.20), $t(214) = 1.48$, p = .14, 95% CI [-0.02, 0.16], d = 0.36. Likewise, in regards to zoning out, those who solved post-incubation did not differ (M = .11, SD = 0.15) from those who never solved (M = .15, SD = 0.17), $t(214) = 1.04$, p = .30, 95% CI [-0.04, 0.12], d = 0.25.

Not surprisingly, cross-experimental analyses of the creativity task confirmed results from Experiments 1 and 2: Regardless of what type of scoring was used, post-incubation divergent thinking was not related to tune outs (average $r(223)$ = -.03, p = .69; top-two $r(223)$ = -.04, p = .56; max-two $r(223)$ = -.02, p = .73) or zone outs (average $r(223)$ = .02, p = .80; top-two $r(223)$ = -.07, p = .29; max-two $r(223)$ = .05, p = .43).

Pre-incubation solvers for the coins and pigpen problems were excluded in the primary analyses, as they were in Experiment 1, but will be reported here for completeness. Once again, subjects who solved the coins task pre-incubation did not differ in WMC z-scores (M = .11, SD = 0.66) from those who never solved (M z-score = -.25, SD = 0.85), $t(92) = -1.56$, $p = .12$, 95% CI [-0.83, 0.10], $d = -0.47$, or those who solved post-incubation (M z-score = -.28, SD = 0.89), $t(35) = -1.46$, $p = .15$, 95% CI [-0.94, 0.15], $d = -0.50$. Likewise, subjects who solved the pigpen problem pre-incubation period did not differ in WMC z-scores (M = -.04, SD = 0.57) from those who never solved (M z-score = -.25, SD = 0.87), $t(106) = -.71$, $p = .48$, 95% CI [-0.80, 0.38], $d = -0.29$, or those who solved post-incubation (M z-score = .23, SD = 0.57), $t(14) = .94$, $p = .36$, 95% CI [-0.35, 0.89], $d = 0.47$.

In addition, subjects who solved the coins before the incubation period did not differ in TUT rates (M = .20, SD = 0.19) from those who never solved (M TUT rate = .28, SD = 0.22), $t(91) = 1.38$, $p = .17$, 95% CI [-0.04, 0.21], $d = 0.42$, or those who solved after the incubation period (M TUT rate = .25, SD = 0.28), $t(34) = .64$, $p = .53$, 95% CI [-0.12, 0.23], $d = 0.23$. Likewise, subjects who solved the pigpen problem before the incubation period did not differ in TUT rates (M = .24, SD = 0.19) from those who never solved (M TUT rate = .29, SD = 0.26), $t(106) = .63$, $p = .53$, 95% CI [-0.12, 0.23], $d = 0.24$, or those who solved after the incubation period (M TUT rate = .14, SD =0.13), $t(14) = -1.10$, $p = .29$, 95% CI [-0.27, 0.09], $d = -0.57$.

Also in line with Experiment 1, pre-incubation creativity was not correlated with WMC z-scores (average scoring: $r(110) = -.03$, $p = .75$; top-two scoring: $r(110) = .07$, $p =$

.45; max-two scoring: *r*(110) = .01, *p* = .93) or TUT rates during the creativity incubation tasks (average scoring: *r*(110) = .07, *p* = .45; top-two scoring: *r*(110) = .00, *p* = 1.00; max-two scoring: *r*(110) = .12, *p* = .21).

With the Raven's problem excluded, the only analytical problem that was used in Experiment 2 was the crime problem. As I found in Experiment 1, there was no statistically significant difference in WMC z-scores between solvers (*M* = -.48, *SD* = 0.93) and non-solvers (*M* = -.15, *SD* =0.82), *t*(113) = 1.64, *p* = .11, 95% CI [-0.07, 0.75], *d* = -0.38.

As noted earlier, the openness subscale comprises six facets (McCrae & Costa, 2010). Of secondary interest was how these facets correlated with insight, creativity, and mind wandering during the reading tasks. In an exploratory analysis, I was especially interested in how the fantasy facet correlated with incubation TUT rates because, with statements like "*I enjoy concentrating on a fantasy or daydream and exploring all its possibilities, letting it grow and develop*" and "*I would have difficulty just letting my mind wander without control or guidance*" (reverse scored), it seemed to have some face validity in measuring a general propensity to mind wander in daily life.

In fact, the fantasy facet was *not* correlated with probed TUT rates collapsed across all incubation reading tasks, *r*(112) = .08, *p* = .42, or tune out rates, *r*(112) = -.06, *p* = .52, but was trending towards a positive correlation with zone out rates, *r*(112) = .18, *p* = .06. Of the remaining facets, none correlated significantly with TUTs or zone outs, but the ideas facet was negatively correlated with tune outs, *r*(112) = -.22, *p* = .02. Consistent with the findings from Baird et al. (2013), the fantasy facet was also positively correlated with post-incubation creativity when the average scoring or max scoring was used (average $r(107) = .25$, $p = .01$; max-two $r(107) = .20$, $p = .04$ and was trending in the same direction when top-two scoring was used, $r(107) = .18$, $p = .07$. Regarding insight, the fantasy facet did predict coins solving, $t(99) = -2.621$, $p = .01$, 95% CI [-1.18, -0.16], $d = -0.66$, with post-incubation solvers scoring higher (*M* z-score = 1.21, *SD* = 0.94) than non-solvers (*M* z-score = .54, *SD* = 1.09), but it did not predict solving in the pigpen problem, $t(104) = 0.943$, $p = .35$, 95% CI [-0.45, 1.28], $d = 0.31$ (post-incubation solvers: *M* z-score = .20, *SD* = 1.57; non-solvers: *M* z-score = .61, *SD* = 1.08).

Discussion

In Experiment 2, I was partly interested in conceptually replicating my results from Experiment 1, but I also added new questions regarding the personality measures. Overall, mind wandering during the attention-demanding tasks decreased from Experiment 1, perhaps because students found the reading tasks to be more interesting and engaging than the SART or n-back tasks used in Experiment 1. I replicated the finding from Experiment 1 that mind wandering during an incubation period is not beneficial for post-incubation insight problem solving or creativity. WMC was again unrelated to post-incubation creativity and post-incubation solving in the coins problem, and although it was not statistically significant (as it was in Experiment 1), it trended towards predicting post-incubation pigpen solving (and this effect was statistically significant in a combined-experiment analysis). Also in line with my Experiment 1 results, I found that the correlation between WMC and TUT rates did not vary according to the context, and concentration did not moderate the relationship between WMC and TUTs.

Cross-experimental analyses of the creativity tasks suggested what both Experiment 1 and Experiment 2 had found separately—neither WMC nor TUT rates predicted post-incubation creativity. Regarding the new personality measures, I found that need for cognition predicted post-incubation performance in the coins insight problem, but not the pigpen problem or creativity tasks. Alternatively, openness to experience generally predicted post-incubation creativity, but did not predict insight. Finally, the fantasy subscale did not correlate with TUT rates, but it did predict post-incubation creativity.

Our inability to replicate the finding from Experiment 1 that post-incubation pigpen solvers have greater WMC than those who never solve is likely due to the small number of post-incubation pigpen solvers in Experiment 2 ($N = 7$), which was half the number from Experiment 1 ($N = 14$). Analyzing the pigpen problem across the two experiments, I found that post-incubation solvers had significantly greater WMC than those who never solved. In addition, post-incubation solvers tended to have lower TUT rates, although this was only trending. Changing the coins problem to a three dimensional format, on the other hand, increased coins solving rates—while there were only 7 post-incubation solvers in Experiment 1, there were 22 post-incubation solvers in Experiment 2—but still, neither WMC nor TUT rates predicted post-incubation solving. Although these two insight problems appear to have different relationships with WMC and TUT rates, this inconsistency is not surprising. Findings from some previous work suggest that the insight construct isn't consistent across problems, and, while I used two visual problems to help unify this construct, it does not appear to have worked. This issue may be a broader problem regarding the field of insight research and will be discussed more in the General Discussion.

Just as I found in Experiment 1, WMC did not predict post-incubation creativity in Experiment 2, regardless of how creativity was scored. This replication goes against the perspective that executive control is beneficial for creativity, but it does add confidence to my initial, null, findings and suggests that my E1 results were *not* due to low power (as does the combined-experiment analysis I reported).

I also replicated my null results from Experiment 1 regarding incubation mind wandering and creativity. Despite efforts to give mind wandering its best possible chance to benefit creativity by changing my incubation task to modified science fiction readings, I found once again that incubation TUT rates did not predict post-incubation creativity.

Once again I tested the prediction that tuning out, but not zoning out, would be beneficial for insight and creativity, and once again I found this not to be the case. Neither tuning out nor zoning out predicted post-incubation performance on the coins task or the divergent thinking tasks, and whereas zoning out was not predictive of pigpen solving either, people who solved the pigpen problem post-incubation actually tuned out *less* than those who never solved. Collapsing across the data sets, cross-experimental analyses confirmed the null results from Experiment 1 that neither tuning out nor zoning out predicted post-incubation performance on either insight or divergent thinking tasks. Overall, it appears that mind wandering is not helpful for insight and creativity, and it makes no difference whether people are aware that they are mind wandering or not. Smallwood, McSpadden, and Schooler (2008) originally suggested that tuning out would be better because people would be aware that their creative insights occurred; however, based on my data, it seems as though people are not having any creative insights during the incubation period, and so it doesn't matter if they are aware of their thoughts or not.

Although initially I had predicted that there would be a negative correlation between WMC and TUTs when there was no reason to mind wander (i.e., during a stand-alone task) and a null or positive correlation between the two when there was (i.e., during an incubation task), results from Experiment 1 suggested a trending pattern in the opposite direction. Because I used two different mind wandering tasks for the two different contexts (i.e., the stand-alone task was always the SART and the incubation task was always an n-back) in Experiment 1, it was difficult to separate the influence of task. In Experiment 2, however, I used a modified reading task for both the stand-alone task and the incubation tasks and I found that the correlation between WMC and TUTs did not vary based on context.

In Experiment 2, I once again found that concentration predicted variation in TUTs, but did not moderate the relationship between WMC and TUTs. Contrary to what previous research into daily life suggests, it does not appear that we are underestimating the relationship between WMC and TUTs in the laboratory, at least based on the current set of experiments.

A new hypothesis in Experiment 2 questioned whether intellectual motivation—as measured by openness and need for cognition—moderated the relationship between

WMC and creativity or WMC and insight. Although I found that openness predicted post-incubation creativity and need for cognition predicted post-incubation coins problem solving, neither measure acted as a moderator. Previous work has suggested that people who score higher on openness to experience and the need for cognition may be more intellectually curious and have a greater motivation to perform well (in addition to actually performing better) on these creativity and insight problems (e.g., Dollinger, 2003; McCrae, 1987). Overall, openness to experience seemed to correlate positively with post-incubation creativity but not post-incubation insight problem solving, and need for cognition was greater for post-incubation coins solvers but did not predict post-incubation pigpen solving or post-incubation creativity. Of primary importance to my hypothesis, however, neither openness nor need for cognition moderated the relationship between WMC and insight or creativity. This suggests that an underlying motivational component does not seem to be driving success on these creativity tasks and insight problems.

CHAPTER IV

GENERAL DISCUSSION

Across two experiments, I tested the influence of individual differences in WMC and mind wandering on insight problem solving and creativity. Overall, the results suggest that WMC positively predicts post-incubation pigpen insight problem solving, but not post-incubation coins solving or divergent thinking. I consistently found that mind wandering during an incubation task was not beneficial for either insight problem solving or creativity. In addition, the correlation between WMC and TUTs did not change with context (i.e., stand-alone vs. incubation task), and it was not moderated by concentration. In Experiment 2, need for cognition positively predicted post-incubation coins solving, but not post-incubation pigpen solving or divergent thinking, and openness to experience predicted post-incubation divergent thinking but not post-incubation insight. Finally, although the fantasy subscale did not correlate with overall TUT rates, it did generally predict post-incubation creativity and post-incubation coins solving (but not pigpen), with higher fantasy scores related to better post-incubation performance.

Our first question was: is executive control helpful for insight and creativity? In Experiment 1 I found that greater WMC predicted post-incubation solving in the pigpen insight problem, but not the coins insight problem. Although changing the coins format to three-dimensional in Experiment 2 did increase the number of solvers (N = 22), WMC still did not predict post-incubation solving in the coins problem. WMC did not significantly predict pigpen post-incubation solving in Experiment 2 either (which saw a decrease in solvers from N = 14 in Experiment 1 to N = 7 in Experiment 2), but the pattern trended in the same direction as what I had found in Experiment 1. In addition, cross-experimental analyses indicated that when the data were collapsed across the two experiments, WMC did predict post-incubation pigpen solving. This finding is not completely unreasonable given the inconsistent findings within the insight literature, suggesting that insight problems do not share much predictive variance, but are rather made up largely of variance specific to each task (Gilhooly & Murphy, 2005). Despite my attempts to find roughly equivalent insight problems, it appears that these tasks have less in common than I had anticipated. Still, previous research on the pigpen problem (e.g., Gilhooly & Murphy, 2005; Gilhooly & Fioratou, 2009) and the coins problem (e.g., Ash & Wiley, 2006) has separately implicated a role for WMC in each of these problems. So why was I able to replicate one but not the other? It could be related to the way the coins problem was displayed. Although both the two dimensional format and the three dimensional format have been used before, previous research has only suggested a role for WMC in the two dimensional format of the problem (Ash & Wiley), and in my current set of experiments, I did not have enough solvers in that format (N = 7) to relate with any other measure. In addition, Ash and Wiley did not present WMC analyses separately for their individual problems; instead, they found that WMC predicted solving in the Many Moves Available category, which included but was not limited to the coins problem in its two dimensional format. For future research interested in normal variation in insight problem solving, it will be extremely important to use many different tasks, rather than one or two.

Across both experiments, and in the cross-experimental analyses, I found that WMC did not predict post-incubation creativity. While some research has suggested that WMC and focused attentional control may be harmful to performance in creativity tasks (Aiello, Jarosz, Cushen, & Wiley, 2012; Jarosz, Colflesh, & Wiley, 2012), these studies have often used the Remote Associates Task, which I would argue does not measure divergent thinking, but rather convergent thinking. While both divergent and convergent thinking seem to be important for creativity, convergent thinking is commonplace in analytical problem solving, but it is divergent thinking that is unique to creativity. The lack of correlation between WMC and creativity is still somewhat surprising, however, given the established relationship between general fluid intelligence and creativity (Gilhooly et al., 2007; Nusbaum & Silvia, 2011; Silvia, 2008). As mentioned previously, general fluid intelligence correlates strongly with WMC and is thought to at least partly reflect a domain general, executive-attention ability. If general fluid intelligence predicts creativity but WMC does not, that would suggest that the relationship between Gf and creativity is not driven by the variance shared between WMC and Gf. Although it is entirely possible, this would be the first time that the shared executive attention was not responsible for explaining variance in higher order cognitive abilities. My null result also seems surprising when I consider WMC's sterling reputation in predicting many higher-order cognitive tasks; however, a couple of newly published studies have also failed to find a relationship between WMC and creativity in zero-order correlations (Lee &

Therriault, 2013; Lin & Lien, 2013). Indubitably, future work will be required to follow up on this result and attempt to separate the variance that is responsible for predicting performance in creativity tasks.

Our second question was: is lack of executive control helpful for insight and creativity? Both experiments here suggest that lack of executive control, as measured by mind wandering, is not helpful for post-incubation insight problem solving or creativity, and may even be detrimental in the pigpen insight problem. This finding contradicts conjecture by Schooler and colleagues that mind wandering should be helpful for creativity, especially during an incubation period (Schooler et al., 2011; Smallwood & Schooler, 2006). Furthermore, my results indirectly contradict recent work by Baird and colleagues (2012) that was published while I was collecting data for Experiment 2.

In their study, Baird et al. (2012) found that the only group of subjects to show improvement in creativity from pre- to post-incubation was the group in the undemanding incubation task condition. Importantly, this same group reported the most mind wandering during incubation according to a retrospective questionnaire. In addition, they found that creativity correlated positively with scores on a general propensity measure of daydreaming in daily life (Imaginal Process Inventory, Daydreaming Frequency subscale; Singer & Antrobus, 1972). Taken together, Baird and colleagues concluded that these results suggest a beneficial role for mind wandering in creative problem solving.

Why did my findings differ from those of Baird et al.? Although the present set of experiments and the study that Baird et al. conducted both use an incubation task inserted into the middle of a creativity task, there are key differences to keep in mind. I measured mind wandering during the incubation tasks using in-the-moment, online thought probes, whereas Baird et al. used a retrospective measure, asking subjects *after* the incubation period to report, on a 1–5 scale, how frequently they were mind wandering during the incubation task. In addition, whereas I found no correlation between incubation TUT rates and post-incubation creativity, Baird and colleagues failed to report this analysis entirely. Instead, they used a two-step, round-about method, whereby they first compared TUT rates across the different incubation tasks and then compared change in pre- to post-incubation creativity across the different incubation tasks, in order to make the claim that TUT rates are related to creativity, when in fact, these are very different conclusions. Using my data, collapsed across both experiments, I ran the analysis that Baird et al. should have ran: I correlated change in pre- to post-incubation divergent thinking performance with incubation TUT rates. But, regardless of whether I used average, top-two, or max-two scoring, change in divergent thinking performance was not correlated with incubation TUT rates (average $r(221) = -.04$, $p = .55$; top-two $r(221) = -.04$, $p = .55$; max-two $r(221) = -.03$, $p = .68$), suggesting that there may have been a reason Baird et al. did not report these analyses. In fact, the only evidence Baird et al. presented which directly implicates mind wandering as beneficial for creativity was their correlation between post-incubation creativity and their broad questionnaire measure of propensity to daydream in daily life, rather than the rates of reported mind wandering during the incubation period. This final point, however, may be the only place where our analyses agree. Although I did not use the same daydreaming questionnaire that they used, the positive correlation that I find between creativity and the fantasy facet of openness does parallel their results.[4]

So why don't probed TUT rates correlate with the broad retrospective measure, and why does only the latter predict creativity? One possibility is that people are able to report their mind wandering in-the-moment, but when they have to aggregate this information and report it in a retrospective questionnaire, they are unable to do so. Instead, perhaps they are reporting self-perception rather than reality in the retrospective questionnaire, and what this correlation is capturing is that people who are more creative tend to think of themselves as daydreamers, waiting for creative inspiration to come. Another possibility is that some people really do mind-wander a lot in daily life, but it doesn't translate into the artificial setting in the lab, and so my in-the-moment probes don't pick up on it.

Our third question was: does the correlation between WMC and TUTs vary by context? Although I did not find evidence of WMC predicting flexibility in mind wandering depending on the context in either of my two experiments here, I may not be able to rule out this hypothesis just yet. Despite efforts to imply that mind wandering may be beneficial during an incubation task by telling subjects they would return to their previous insight or creativity task immediately after the incubation period, it is possible

---

[4] In a third experiment that will be included for publication, I modified my current design from Experiments 1 and 2 to include Baird et al.'s shorter incubation task along with the Imaginal Process Inventory for daydreaming with the intention of providing conditions that would give mind wandering its best possible chance to benefit creativity. I also included the Imaginal Process Inventory for mind wandering and kept my online thought probes to measure in-the-moment mind wandering during the incubation period.

that this may not have been obvious enough of a cue. Future work may consider a more direct approach whereby subjects are either told that mind wandering could be beneficial during an incubation period or they are simply instructed to mind wander during the incubation task.

Our fourth question was: does concentration moderate the relationship between WMC and TUTs? Past research into daily life has found an association between WMC and TUTs, but only when concentration is high (Kane et al., 2007). Since previous lab research has not taken concentration into account, I tested the possibility that the correlation between WMC and TUTs has been underestimated. In both experiments, however, I found this not to be the case. Although I found considerable variability in my concentration measure across these two studies, it is likely that concentration on daily life tasks is inherently more variable than that required across similar lab tasks.

Our fifth question, which was only tested in Experiment 2, was: does intellectual motivation moderate the relationship between WMC and insight or creativity? Using the openness subscale of the NEO along with the need for cognition questionnaire to ascertain estimates of effort and how highly subjects valued the goal of the task, I hypothesized that these measures would moderate the association between WMC and insight or creativity. Given my failure to find an association between WMC and creativity, I was particularly interested in whether a pattern would emerge once intellectual motivation was taken into account—perhaps WMC would predict creativity when intellectual motivation was high. Although I found that openness predicted creativity and need for cognition predicted coins problem solving, neither measure moderated the relationship between WMC and creativity or insight. The inconsistency in how need for cognition related to the insight problems provides further evidence that I was not tapping into a unified insight construct with the coins and pigpen problem.

In conclusion, I demonstrated across two experiments that mind wandering is not beneficial for creativity or insight problem solving, and could in fact be harmful. Alternatively, WMC did predict post-incubation pigpen solving, but not post-incubation coins solving or post-incubation creativity. Although my results were inconsistent between the two insight tasks, it is safe to conclude that WMC is not harmful to insight problem solving and that mind wandering certainly is not helpful. Regarding creativity, there appears to be no association with WMC or in-the-moment mind wandering during an incubation period—null results which were replicated across two experiments with considerable sample sizes ($N > 120$ in each). In our second experiment, however, the fantasy facet of openness *did* correlate positively with creativity. Regarding task context, the correlation between WMC and TUTs did not fluctuate based on the potential for benefit (i.e., during an incubation task). Finally, my measures of intellectual motivation—openness to experience and need for cognition—did not interact with WMC to predict performance in insight or creativity.

REFERENCES

Ash, I. K., & Wiley, J. (2006). The nature of restructuring in insight: An individual-differences approach. *Psychonomic Bulletin & Review*, *13*(1), 66–73.

Baars, B. J. (2010). Spontaneous repetitive thoughts can be adaptive: Postscript on “mind wandering.” *Psychological Bulletin*, *136*(2), 208–210.

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W. Y., Franklin, M. S., & Schooler, J. W. (2012). Inspired by Distraction: Mind Wandering Facilitates Creative Incubation. *Psychological Science*, *23*(10), 1117–1122.

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition*, *20*(4), 1604–1611.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology: Monograph*, *80*(3), 1–46.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48,* 306–307.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Chein, J. M., Weisberg, R. W., Streeter, N. L., & Kwok, S. (2010). Working memory and insight in the nine-dot problem. *Memory & Cognition*, *38*(7), 883–892.

Chronicle, E. P., MacGregor, J. N., & Ormerod, T. C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 14–27.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Colflesh, G. J. H., & Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review*, *14*(4), 699–703.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*(2), 331–335.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 125–155). Cambridge, MA: MIT Press.

Dijksterhuis, A., & Meurs, T. (2006). Where creativity resides: The generative power of unconscious thought. *Consciousness and Cognition*, *15*, 135–146.

Dollinger, S. J. (2003). Need for uniqueness, need for cognition, and creativity. *The Journal of Creative Behavior, 37,* 99–116.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *The Psychology of Learning and Motivation* (pp. 145–199).

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128,* 309–331.

Fleck, J. I. (2008). Working memory demands in insight versus analytic problem solving. *European Journal of Cognitive Psychology*, *20*(1), 139–176.

Gilhooly, K. J., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking & Reasoning*, *15*(4), 355–376.

Gilhooly, K.J., Fioratou, E., Anthony, S.H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98,* 611–625.

Gilhooly, K. J., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking & Reasoning*, *11*(3), 279–302.

Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill.

Hambrick, D.Z., & Engle, R.W. (2003). The role of working memory in problem solving. In J.E. Davidson & R.J. Sternberg (Eds.), *The psychology of problem solving* (pp. 176–206). Cambridge: Cambridge University Press.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26,* 107–129.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., . . . Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *Public Library of Science Biology, 4*, 1–23.

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

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). NY: Oxford.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66–71.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*(3), 374–419.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1534–1555.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence, 14*, 389–433.

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, *41*, 306–320.

Levinson, D. B., Smallwood, J., & Davidson, R. J. (2012). The persistence of thought: Evidence for a role of working memory in the maintenance of task-unrelated thinking. *Psychological Science*, *23*(4), 375–380.

Lin, W. L., & Lien, Y. W. (2013). The different role of working memory in open-ended versus closed-ended creative problem solving: A dual-process theory account. *Creativity Research Journal*, *25*(1), 85–96.

McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, *52*(6), 1258–1265.

McCrae, R. R., & Costa, P. T., Jr. (2010). *NEO Inventories for the NEO Personality Inventory-3 (NEO-PI-3), NEO Five-Factor Inventory-3 (NEO-FFI-3), and NEO Personality Inventory–Revised (NEO PI-R) professional manual*. Lutz, FL: PAR.

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.

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

McVay, J. C., & Kane, M. J. (2012a). 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.

McVay, J. C. & Kane, M. J. (2012b). Why does working memory capacity predict variation in reading comprehension? On the influence on mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*(2), 302–320.

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*, 136–147.

Mullis, K. (1998). *Dancing naked in the mind field*. New York, NY: Pantheon Books.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence*, *39*(1), 36–45.

Oberauer, K., Schulze, R., Wilhelm, O., & Süβ, H. M. (2005). Working memory and intelligence--their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61–65.

Ohlsson, S. (1992). Information processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking*. London: Harvester-Wheatsheaf.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Cognition*, *28*(4), 791–799.

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary scales: Section 1. General overview.* San Antonio: Harcourt Assessment.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211–227.

Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology*, *84*, 171–199.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*(2), 166–183.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.), *Thinking and Seeing: Visual Metacognition in Adults and Children* (pp. 203–226). Cambridge, MA: MIT Press.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The Nature of Insight* (pp. 65–124). Cambridge, MA: MIT Press.

Silvia, P. J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, *44*(4), 1012–1021.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 68–85.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Dialogue: The Official Newsletter of the Society for Personality and Social Psychology*, *26*, 4–7.

Singer, J. L., & Antrobus, J. S. (1972). Daydreaming, imaginal processes, and personality: A normative study. In P. W. Sheehan (Ed.) *The Function and Nature of Imagery* (pp. 175–202). New York: Academic Press, Inc.

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94–120.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, *14*(3), 527–533.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: the curious incident of the wandering mind. *Memory & Cognition*, *36*(6), 1144–1450.

Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition, 18,* 118–125.

Smallwood, J., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic sustained attention to response task. *Consciousness and Cognition*, *15*, 218–231.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958.

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology*, *104*(1), 61–87.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505.

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

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289–335.

Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The Nature of Insight* (pp. 157–196). Cambridge, MA: MIT Press.

Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of “fixation” in the solution of several “insight” problems. *Journal of Experimental Psychology: General*, *110*(2), 169–192.

White, H., & Shah, P. (2006). Uninhibited imaginations: Creativity in adults with Attention-Deficit/Hyperactivity Disorder. *Personality and Individual Differences*, *40*(6), 1121–1131.

Wieth, M. B., & Zacks, R. T. (2011). Time of day effects on problem solving: When the non-optimal is optimal. *Thinking & Reasoning*, *17*(4), 387–401.

APPENDIX A

N-BACK STIMULI FOR EXPERIMENT 1 MIND WANDERING TASKS

Countries	Body Parts	Instruments	Colors
France	Heart	Oboe	Green
China	Foot	Violin	Orange
Spain	Head	Organ	Black
Italy	Hand	Banjo	Pink
Russia	Brain	Piano	White
Brazil	Nose	Cymbal	Brown
Mexico	Arms	Tuba	Gray
India	Legs	Flute	Red
Greece	Neck	Drum	Gold
Sweden	Ear	Harp	Yellow
Japan	Eye	Sax	Purple
Canada	Mouth	Guitar	Blue

APPENDIX B

EXPERIMENT 1 MEAN ACCURACY AND REACTION TIMES BY TASK

	Accuracy		RT	
	M	SD	M	SD
SART Target Trials	.420	.222	369.469	77.352
SART Non-Target Trials	.933	.049	446.479	89.661
Coins N-back Target Trials	.772	.196	850.984	230.546
Coins N-back Non-Target Trials	.973	.046	1095.834	574.543
Coins N-back Lure Trials	.870	.108	981.211	298.798
Pigpen N-back Target Trials	.793	.186	843.432	229.570
Pigpen N-back Non-Target Trials	.977	.045	1010.872	464.521
Pigpen N-back Lure Trials	.877	.100	1058.591	321.313
Brick N-back Target Trials	.768	.215	840.427	229.886
Brick N-back Non-Target Trials	.970	.048	909.151	465.074
Brick N-back Lure Trials	.852	.112	1009.053	280.958
Knife N-back Target Trials	.765	.195	844.443	233.704
Knife N-back Non-Target Trials	.971	.045	984.665	519.659
Knife N-back Lure Trials	.876	.097	1053.134	367.940

Note. SART = sustained attention response task

APPENDIX C

EXPERIMENT 2 MEAN ACCURACY AND REACTION TIMES BY TASK

	Accuracy		RT	
	M	SD	M	SD
Stand-alone Reading Task Target Trials	.440	.169	492.416	78.146
Stand-alone Reading Task Non-Target Trials	.981	.022	514.482	111.037
Coins Reading Task Target Trials	.521	.168	473.221	91.447
Coins Reading Task Non-Target Trials	.983	.024	487.750	103.516
Pigpen Reading Task Target Trials	.497	.179	480.617	99.992
Pigpen Reading Task Non-Target Trials	.983	.020	489.136	116.983
Brick Reading Task Target Trials	.484	.185	484.864	93.675
Brick Reading Task Non-Target Trials	.985	.014	463.303	111.890
Knife Reading Task Target Trials	.480	.172	484.397	97.585
Knife Reading Task Non-Target Trials	.983	.026	482.256	117.217
Stand-alone Reading Comprehension Questions	4.388	.892	–	–
Coins Reading Comprehension Questions	4.548	.740	–	–
Pigpen Reading Comprehension Questions	4.362	.973	–	–
Brick Reading Comprehension Questions	4.422	.925	–	–
Knife Reading Comprehension Questions	4.522	.809	–	–

Note. Reading comprehension questions were scored out of 5.