

## Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought

By: Bridget A. Smeekens, [Michael J. Kane](#)

Smeekens, B.A., & Kane, M.J. (2016). Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought. *Psychology of Aesthetics, Creativity, and the Arts*, 10(4), 389-415.

©American Psychological Association, 2016. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at: <http://dx.doi.org/10.1037/aca0000046>

\*\*\*© APA. Reprinted with permission. No further reproduction is authorized without written permission from APA. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. \*\*\*

### **Abstract:**

Should executive control, as indicated by working memory capacity (WMC) and mind-wandering propensity, help or hinder creativity? Sustained and focused attention should help guide a selective search of solution-relevant information in memory and help inhibit uncreative, yet accessible, ideas. However, unfocused attention and daydreaming should allow mental access to more loosely relevant concepts, remotely linked to commonplace solutions. Three individual-differences studies inserted incubation periods into 1 or 2 divergent thinking tasks and tested whether WMC (assessed by complex span tasks) and incubation-period mind wandering (assessed as probed reports of task-unrelated thought [TUT]) predicted postincubation performance. Retrospective self-reports of Openness (Experiment 2) and mind-wandering and daydreaming propensity (Experiment 3) complemented our thought-probe assessments of TUT. WMC did not correlate with creativity in divergent thinking, whereas only the questionnaire measure of daydreaming, but not probed thought reports, weakly predicted creativity; the fact that in-the-moment TUTs did not correlate with divergent creativity is especially problematic for claims that mind-wandering processes contribute to creative cognition. Moreover, the fact that WMC tends to strongly predict analytical problem solving and reasoning, but may not correlate with divergent thinking, provides a useful boundary condition for defining WMC's nomological net. On balance, our data provide no support for either benefits or costs of executive control for at least 1 component of creativity.

**Keyword:** divergent thinking | executive attention | individual differences | mind wandering | working memory

### **Article:**

What's the best way to creatively generate a new idea? Should we cognitively buckle down, as when wrestling with a complex syllogism in logic class, and attempt to screen distractions from external stimuli and task-irrelevant thoughts? Or, might we be better off loosening up, allowing

the mind to roam freely to sights, sounds, and ideas that seem only remotely connected to the task at hand? To determine whether creativity comes to those who work or to those who wander, we heeded Underwood's (1975) advice that individual differences provide a critical test-bed for theorizing. Specifically, in three studies we examine the extent to which two indices of executive control—working memory capacity and propensity for mind wandering—predict normal variation in divergent thinking.

The literature already indicates that individual differences in working memory capacity (WMC) predict important cognitive abilities and outcomes, such as reasoning (e.g., Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & Süß, 2005), language comprehension (e.g., Daneman & Merikle, 1996; Just & Carpenter, 1992), multitasking (Hambrick, Oswald, Darowski, Rench, & Brou, 2010; König, Bühner, & Mürling, 2005), learning (e.g., Engel de Abreu & Gathercole, 2012; Shute, 1991), and academic performance (e.g., Cowan et al., 2005; Gathercole, Pickering, Knight, & Stegmann, 2004). The breadth of these associations, along with research showing that WMC variation also correlates with performance on “simpler” attention-control tasks (e.g., Conway, Cowan, & Bunting, 2001; Kane & Engle, 2003; Unsworth & Spillers, 2010), has led some theorists to propose that domain general executive-control mechanisms contribute significantly to WMC and its predictive power (e.g., Braver, Gray, & Burgess, 2007; Engle & Kane, 2004; Hasher, Lustig, & Zacks, 2007; Unsworth & Engle, 2007).

But that predictive power has limits. In domains of lower-level cognition, the executive processes related to WMC appear important to restraining habitual but inappropriate responses (e.g., Kane, Bleckley, Conway, & Engle, 2001; Hutchison, 2011; Unsworth, Schrock, & Engle, 2004), constraining attentional focus amid distractors (e.g., Ahmed & de Fockert, 2012; Heitz & Engle, 2007; Redick & Engle, 2006), and sustaining conscious focus to ongoing tasks (e.g., McVay & Kane, 2009, 2012a; Unsworth, Redick, Lakey, & Young, 2010). They do not, however, seem to contribute to guiding visual search through large stimulus arrays (e.g., Kane, Poole, Tuholski, & Engle, 2006; Poole & Kane, 2009; Sobel, Gerrie, Poole, & Kane, 2007) or switching rapidly between competing task sets (e.g., Kane, Conway, Hambrick, & Engle, 2007; Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000; Oberauer, Süß, Wilhelm, & Wittmann, 2003). Dissociations like these suggest boundary conditions to the WMC-executive association, which should benefit theory by specifying the nomological net (Cronbach & Meehl, 1955; Meehl, 1978) around the still-vague executive attention construct.

Here we explore whether similar dissociations arise in a domain of higher-order cognition, suggesting that WMC is important to some complex cognitive abilities but not to others; indeed, we will test whether executive control may actually be counterproductive in some task contexts. On one hand, many complex cognitive abilities, such as reasoning and analytical problem solving, rely on executive-control processes to keep task goals, rules, and partial solutions accessible and shielded from distraction, as well as to guide strategic memory search (e.g., Cho, Holyoak, & Cannon, 2007; Hambrick & Engle, 2003; Kaplan & Simon, 1990; Wiley & Jarosz, 2012). On the other hand, insightful problem solving and divergent thinking may not. The literature is actually mixed regarding the role that executive processes play in these creative domains—some argue that executive control is helpful for insight and divergent thinking

(e.g., Beaty & Silvia, 2012; Gilhooly & Fioratou, 2009; Gilhooly, Fioratou, Anthony, & Wynn, 2007; Nusbaum & Silvia, 2011), whereas others maintain that a lack of cognitive control is ideal (Baird et al., 2012; Chrysikou & Thompson-Schill, 2011; Limb & Braun, 2008; Reverberi, Toraldo, D'Agostini, & Skrap, 2005).

Our three experiments will assess the influence of executive control on creativity in two ways: (a) by correlating performance on divergent thinking tasks with WMC, and (b) by borrowing a design from the incubation literature to measure propensity for mind wandering, or task-unrelated thoughts (TUTs), during breaks within divergent thinking tasks, to see whether frequent TUT experiences are associated with more or less creative divergent thinking. Although correlation does not necessarily imply causation, causation does imply statistical association, and so current theoretical perspectives that claim cognitive control benefits creative thinking also predict that people with better control (i.e., greater WMC) should be the most creative, whereas theories that claim a lack of cognitive control benefits creative thinking also predict that people with worse control (i.e., greater mind-wandering propensity) should be the most creative. Strictly speaking, the association implied by causation need not be linear, but current theoretical claims imply linearity. Our study's correlational methods put these theoretical claims to a critical test (à la Underwood, 1975).

### **Potential Benefits of Executive Control to Creativity**

Generating novel and useful interpretations of old items and ideas may be hindered by one's knowledge of norms and commonplace uses (e.g., Gilhooly et al., 2007), and so innovative and creative ideas arise only after abandoning these constraints. For example, in the alternative uses task, a common laboratory measure of divergent thinking (Guilford, 1967), subjects initially output "creative" uses for common objects that were simply retrieved from memory (Gilhooly et al., 2007). These mundane responses are easily accessible and so people who stop thinking about them and switch to a more effective strategy will generate more creative responses (Gilhooly et al., 2007). Indeed, when subjects must subsequently distinguish their responses that were retrieved from memory from those generated on the spot, their executive-control abilities (indexed by verbal fluency) predicted on-the-spot responding but not retrieved-from-memory responding (Gilhooly et al., 2007). Thus, executive control might facilitate creative thinking.

Few studies have assessed the relation between WMC and divergent thinking, and they have yielded mixed results (Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; De Dreu et al., 2012; Lee & Therriault, 2013; Lin & Lien, 2013). However, a growing literature has suggested a contribution of general fluid intelligence (Gf), the domain-general ability to reason through novel analytical problems (Carroll, 1993; Horn & Cattell, 1967). Gf correlates strongly with WMC ( $\approx .70$ – $.80$  at the latent-variable level; Kane, Hambrick, & Conway, 2005; Oberauer et al., 2005) and moderately with creative responding in divergent thinking, alternative-uses, tasks ( $\approx .35$ – $.45$  at the latent level; e.g., Gilhooly et al., 2007; Silvia, 2008a; see also Batey, Furnham, & Safiullina, 2010), suggesting that the shared executive-related variance between WMC and Gf might predict divergent thinking. Moreover, Nusbaum and Silvia's (2011, Study 1) latent-variable analyses indicated that the association between Gf and divergent thinking was mediated by the executive process of switching between categories of alternate uses (i.e., total number of

broad categories of uses) but not clustering (i.e., number of uses produced in each category), which may involve successful inhibitory control (e.g., Gilhooly et al., 2007). Their Study 2 provided half the subjects with a helpful strategy for producing creative uses in the divergent thinking task, specifically to disassemble the object and reassemble the parts. Strategy instruction amplified, rather than reduced, the Gf-divergent thinking correlation, indicating that the executive abilities associated with Gf help maintain a strategy in memory and apply it effectively.

### **Potential Costs of Executive Control to Creativity**

The evidence reviewed so far indicates that executive control is generally helpful to creative cognition. But other findings, from the Remote Associates Test (RAT) and from the mind-wandering literature, suggest otherwise. The RAT presents subjects with three seemingly unrelated words that are, in fact, associated with a common word or concept (e.g., PAINT—DOLL—CAT: “*house*”; Mednick & Mednick, 1967). Solving such problems requires thinking divergently, to explore the semantic space for each word, and then convergently, to derive the common associate, and is sometimes accompanied by a subjective “aha” feeling of insight (e.g., Bowden & Jung-Beeman, 2003; Jung-Beeman et al., 2004). Although solving RAT problems is sometimes positively correlated with WMC (De Dreu et al., 2012; Kane et al., 2004; Ricks, Turley-Ames, & Wiley, 2007), it is also sometimes facilitated by thwarting cognitive control via alcohol intoxication or instruction. Jarosz, Colflesh, and Wiley (2012) equated two groups of subjects on WMC using a verbal-symbolic 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 the span task again as well as the RAT. The sober group showed positive practice effects on the WMC task, but the intoxicated group did not, suggesting decreased executive control in those who consumed alcohol. At the same time, the intoxicated group solved more RAT problems, did so faster, and experienced more subjective insight than did controls. Strategic suppression of control may also aid RAT performance: Aiello, Jarosz, Cushen, and Wiley (2012) instructed half their subjects to “use your gut” in solving RAT problems and they performed better than did uninstructed control subjects. The authors suggest that the “use your gut” instruction reduced subjects’ reliance on attentional control mechanisms to solve the problems.

Propensity for mind wandering during ongoing tasks also seems to be negatively associated with control abilities and positively associated with creative thought. During challenging tasks, at least, people with lower WMC report more TUTs than do those with higher WMC (Kane et al., 2007; McVay & Kane, 2009, 2012a, 2012b; Unsworth, Brewer, & Spillers, 2012; Unsworth & McMillan, 2013; Unsworth, McMillan, Brewer, & Spillers, 2012; for a meta-analytic review, see Randall, Oswald, & Beier, 2014). Some theorists argue that mind wandering is automatically triggered by environmental and mental cues to personal goals and concerns, and to represent, in part, a failure of the executive control system to block interference from TUTs (Kane & McVay, 2012; McVay & Kane, 2010; see also Smallwood, 2013). Based on the typically negative consequences that are associated with attention-control failure, then, one might predict that mind

wandering would be harmful to performance in complex and demanding tasks such as creative problem solving.

Alternatively, some have suggested that mindwandering experiences, in which we mentally escape the confines of the present, may be beneficial to personal goal striving (e.g., Klinger, 1971, 2013; McMillan, Kaufman, & Singer, 2013; Singer, 1966) and to creativity (Baars, 2010; Schooler et al., 2011; Smallwood & Schooler, 2006). Informal support comes from countless anecdotes that illustrate fruitful mind wandering during an incubation period, from Archimedes' apocryphal "eureka" moment in the bath, to Nobel laureate Kary Mullis's insight into amplifying DNA sequences while driving along a California freeway (Mullis, 1998). Direct empirical support, however, is harder to come by. A small study by Singer and Schonbar (1961) found that graduate students scoring higher on a retrospective daydreaming questionnaire wrote more creative accounts of a daydream and a more creative fictional story. In contrast, a large-scale study of undergraduates (Singer & Antrobus, 1963) found no associations between a battery of divergent thinking tasks and various daydreaming-questionnaire factors; in exploratory factor analyses, the only daydream-related factor that also included divergent-thinking scores reflected uncontrolled mind wandering and distractibility and, here, the loading for divergent thinking was *negative*, indicating poorer creativity in the most frequent mind-wanderers.

Indirect empirical evidence from the incubation literature, however, supports a positive role for mind wandering in creative thinking. Dozens of studies have evaluated the effects of various incubation tasks (i.e., taking a break from an initial task to work on an unrelated task) on insight problem solving and divergent thinking; although some found that an incubation period benefits subsequent performance (known as the "incubation effect"), others have not. A meta-analysis categorizing studies by problem type (e.g., divergent thinking, visual insight, or verbal insight problems) and incubation-task demands (e.g., rest, low cognitive demand, high cognitive demand) found a modest incubation benefit ( $M$  effect size = 0.29; 95% CI [0.21, 0.39]; Sio & Ormerod, 2009). Incubation periods had positive effects on all three task types ( $M$  effect sizes = .22–.29), but moderation analyses specified that low-demand incubation tasks yielded better subsequent performance on divergent thinking and verbal insight tasks than did high-demand tasks (but limited data-points for divergent thinking prevented strong conclusions about cognitive demand).

Schooler et al. (2011) thus argued that incubation tasks with low cognitive demands should allow abundant opportunities for TUTs, whereas tasks with high demands would not (e.g., Antrobus, Singer, & Greenberg, 1966; Teasdale et al., 1995; Teasdale, Proctor, Lloyd, & Baddeley, 1993), and suggested that mind wandering might thus contribute to creative thinking. Smallwood and Schooler (2006, p. 956) also drew 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)." While we collected data for our second experiment here, Baird et al. (2012) reported suggestive findings that incubation-period mind wandering increased divergent creativity. Subjects completed one of four incubation tasks—an undemanding task, a demanding task, unstructured rest, or no break—midway through a set of divergent thinking tasks to generate novel uses for

everyday objects. Following incubation, subjects retrospectively rated their TUT frequency and then returned to the divergent thinking tasks. Baird et al. did not report the correlation between divergent thinking scores and self-reported TUT rate during incubation, which was actually critical to their theoretical claim that mind-wandering processes affected creativity. They did find, however, that the undemanding incubation condition produced both the most mind wandering and the most improvement in divergent creativity scores from pre- to postincubation (i.e., there was a *correlation* between the experimental effect on mind wandering and the experimental effect on divergent creativity; a causal path from mind wandering to creativity should not be implied). In contrast to findings from Singer and Antrobus (1963), postincubation divergent creativity scores correlated positively ( $r \approx .20$ ) with scores on a general retrospective questionnaire of daydreaming frequency in daily life (Singer & Antrobus, 1972), collapsed across all experimental groups. Mind wandering and creativity may thus appear to be linked in some way (but see Singer & Antrobus, 1963), but there is currently no direct evidence for the causal claim that mind wandering facilitates creativity, or its individual-differences corollary that people who experience more TUTs during a particular incubation task will thus perform more creatively than will those who experience less off-task thinking.

## **The Present Work**

We suggest that the literature—including studies of WMC and mind wandering—is mixed regarding the benefits versus costs of executive-control variation to individual differences in creative cognition. In the current investigation, then, we focused on multiple measures of WMC and incubation-period mind-wandering propensity as indicators of executive control (and its failure), and multiple measures of divergent thinking as markers of creative cognition. For WMC, we created a composite score for each subject from two complex span tasks, operation span and symmetry span (Conway et al., 2005). For creative cognition, we had hoped to assess both convergent creativity, where people must derive a particular correct answer or product, and divergent creativity, where there is not only one answer or product. We measured convergent creativity with two visual insight problems (in both Experiments 1 and 2), and we measured divergent creativity by combining three raters' judgments of creative responding across two divergent thinking tasks that asked subjects to generate clever, original, unusual, uncommon, and creative uses for an object (Experiments 1–3). Unfortunately, very few subjects solved the insight problems in either experiment, and the solution rates for each problem were highly inconsistent across experiments. We therefore were unable to draw conclusions from these tasks and will not discuss them further (beyond what is required to make the overall study procedures clear). We measured mind-wandering propensity with unpredictable thought probes that were embedded into ongoing tasks — during a baseline, stand-alone task and, of most importance, during the incubation tasks inserted into the insight problems and divergent thinking tasks.

### **Experiment 1**

To test whether incubation improves divergent thinking because it promotes mind wandering, we chose two tasks that we knew provided a supportive environment for TUTs: the Sustained Attention to Response Task (SART), for the stand-alone assessment of TUTs, and the *n*-back task, for the incubation-task assessment of TUTs. Previous work has consistently shown that

subjects mind wander approximately 30% to 50% of the time while working on the SART (e.g., McVay & Kane, 2009, 2012a; McVay, Meier, Touron, & Kane, 2013); the *n*-back has been used successfully as an incubation task (Dijksterhuis & Meurs, 2006) and induces comparable TUT rates to the SART (McVay et al., 2013).

## **Method**

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in these studies (Simmons, Nelson, & Simonsohn, 2012).

## **Subjects**

Undergraduates enrolled in introductory psychology courses at the University of North Carolina at Greensboro (UNCG) earned partial credit toward a course requirement for participating in two 120 min sessions. Using two complete semesters as our data-collection stopping rule, 173 students completed the first session and 142 returned to complete the second in that time. To minimize attrition, we gave most credits upon completion of Session 2. Unless stated otherwise, data analyses included only 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 (Redick et al., 2012; Unsworth, Heitz, Schrock, & Engle, 2005; Unsworth, Redick, Heitz, Broadway, & Engle, 2009). Both tasks—operation span and symmetry span—required subjects to perform a processing task while simultaneously remembering short lists of unrelated items in serial order. Each task began with three practice blocks: The first allowed subjects to practice the memory component for four trials (two each of set sizes two and three), the second provided practice on the processing component for 15 trials, and the third combined both the memory and processing components for 8 trials. For critical test trials, response deadlines for the processing task ensured that subjects did not pause to rehearse the memory items; deadlines were calculated individually for each subject using the response times from the processing-component practice block ( $M + 2.5 SDs$ ). If subjects exceeded the deadline on any trial, the task moved on and counted that item as an error.

### *Operation span*

The processing component of this task required subjects to solve a math problem [for example,  $(9 \div 3) + 2 = ?$ ] and then evaluate the solution (e.g., 6) presented on the next screen by clicking the computer mouse on the word TRUE or FALSE; each equation was followed by one of 12 possible letters to remember F, K, P, S, H, L, Q, T, J, N, R, and Y). After 3 to 7 equation-letter pairs, subjects saw the 12 letters onscreen and used the mouse to select the targets in their serial order. A total of 15 trials presented 3 sets of each size (3–7) in random order.

### *Symmetry span*

Subjects evaluated the vertical symmetry of a black-and-white  $8 \times 8$  grid pattern while remembering the location of a subsequently presented red square within a  $4 \times 4$  matrix. After 2 to 5 symmetry-square pairs, subjects saw an empty  $4 \times 4$  matrix and used the computer mouse to

recall the locations of the red squares in serial order. A total of 12 trials presented 3 sets of each size (2–5) in random order.

### **Mind-wandering assessment: Ongoing tasks and thought probes**

Mind wandering was measured by randomly probing subjects about their immediately preceding thoughts during 20-min cognitive tasks. The SART served as a stand-alone measure and four versions of the *n*-back served as incubation tasks.

#### *SART*

In this go/no-go task, subjects viewed a sequence of words, presented one at a time, and decided whether each was an animal or a food (McVay & Kane, 2009). Each word appeared for 300 ms, followed by a 900-ms mask. If the word was an animal, subjects pressed the space bar as quickly as possible (89% of the 900 trials). If the word was a food, subjects withheld response and pressed no key (11% of trials). Following 60% of the critical, no-go stimuli, a thought probe asked subjects to characterize the content of their thoughts in the moment preceding the probe (see below for details).

#### *n-back*

Subjects decided whether each sequentially presented word matched the one presented two words back. If the words matched (e.g., green, blue, *green*), subjects pressed the space bar as quickly as possible (25% of the 336 trials). If the words didn't match, subjects withheld response (75% of trials); a subset of nontarget trials (21%) were lures that presented either 1-back matches (e.g., blue, green, *green*) or 3-back matches (e.g., green, blue, purple, *green*). Each word appeared for 500 ms, followed by a 2500-ms fixation cross. Three thought probes appeared unpredictably during each of the seven task blocks, two of which appeared immediately after a 2-back target. Four versions of the *n*-back functioned as the incubation-period task for the two insight problems and two divergent thinking tasks, each presenting different stimulus categories (from Battig & Montague, 1969; Van Overschelde, Rawson, & Dunlosky, 2004): colors, countries, body parts, and musical instruments (see Appendix B for item lists).

#### *Mind-wandering probes*

Each probe screen asked, “What were you just thinking about?” Subjects’ reported their thoughts via key-press to these on-screen choices: (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). We explained “tuning out” as when a person is fully aware that she is mind-wandering, and “zoning out” as when she doesn’t realize that she is mind-wandering until something in the environment (like the thought probe) interrupts her, or she catches herself. We took instructions to subjects from one of several published studies that distinguished tune-outs from zone-outs (Smallwood, McSpadden, & Schooler, 2007, p. 533), with only slight modifications and with some text varying according to the ongoing task in which the probes were embedded:

During this experiment you will be asked at various points whether your attention is firmly directed toward the task, or alternatively you may be aware of other things besides the task.



Occasionally you may find as you are performing the task that you are thinking about something completely unrelated to what you are doing; 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 do the task. This is what we refer to as “tuning out”— that is, when your mind wanders and you know it all along.

**ZONING OUT.** Other times when your mind wanders, you do not realize that your thoughts have drifted away from the task until you catch yourself. This is what we refer to as “zoning out”— that is, when your mind wanders, but you do not realize this until you catch it.

Schooler has informally speculated (quoted in Glausiusz, 2011; Tierney, 2010) that distinguishing tune-outs from zone-outs may be important because creative insights while zoning out may go unnoticed. It is not clear to us, however, whether the tune-out/zone-out distinction— as understood by research subjects —captures anything more than a continuum from less to more off-task. Therefore, our analyses scored responses of either “3” (tuning out) or “4” (zoning out) as a TUT, or mind-wandering experience; we note, however, that in exploratory analyses we assessed tune-outs and zone-outs separately, but in no case did the findings from these differ from each other or from the overall TUT pattern, so we do not report them here. We considered responses of “2”, reflecting thoughts about one’s own performance, as “task related interference” (TRI; Smallwood, Riby, Heim, & Davies, 2006), which does not represent either fully on-task or off-task thought (McVay & Kane, 2009, 2012a, 2012b; McVay et al., 2013).

### **Divergent creativity: Alternative uses tasks**

Two versions of the “alternative uses” task (e.g., Guilford, 1967) assessed individual differences in divergent creativity. In both, subjects generated as many creative uses for an everyday object (a brick or a knife) as they could. After 5 min on the task, subjects switched to the incubation *n*-back task for approximately 20 min, and then switched back to the same alternative uses task for another 5 min (see Figure 1). To best assess *creative* thinking, our 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. Without such “be creative” instructions, divergent thinking tests resemble simple verbal-fluency tasks more than creativity tasks (Nusbaum, Silvia, & Beaty, 2014; see also Harrington, 1975; Silvia, 2008a). Below are the relevant verbatim instructions for the Knife task:

In this task, we are interested in how creative and original you can be in coming up with new ideas - here, in thinking up unique and clever ways to use an everyday object. Specifically, we would like you to generate as many original, clever, and creative uses for a knife that you can think of.

Certainly there are many common and everyday ways to use a knife. But for this task, we want you to list all of the unusual and uncommon uses that you can invent or think of. Try to think creatively, and try to come up with clever uses for a knife that are not like any uses that you’ve

ever seen or heard of before. Your goal is to try to develop such original and clever uses for a knife that few other UNCG students will come up with the same ideas as you.

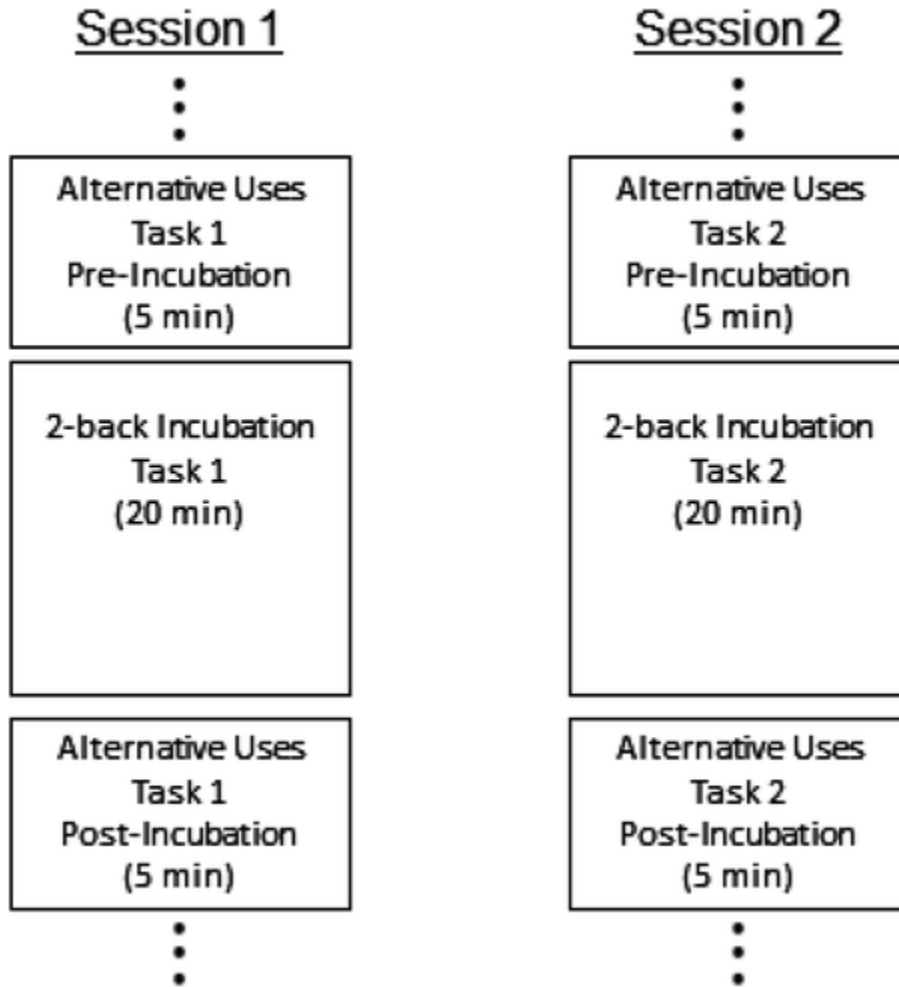


Figure 1. Incubation design for alternative uses task in Experiment 1. One alternative uses task was completed in Session 1 (e.g., alternative uses for a brick) and another in Session 2 (e.g., alternative uses for a knife) counterbalanced across subjects.

After each 5-min period, subjects saw a list of their responses from that period on-screen and retyped what they thought were their two best answers; after the second period they also subsequently chose their best two answers from the full 10 min. Asking subjects for their best responses allowed us to use “top-two” scoring, in addition to a “total” score that averaged across all of their responses. We provide more information about scoring below.

### General procedure

All subjects were tested individually by an undergraduate or graduate experimenter who remained in the testing room throughout the session. Dell desktop computers, with 17-in. LCD or CRT monitors and running E-Prime 1.2 software, presented all task stimuli and collected all

responses. In the first session, subjects completed the SART, one of the alternative uses tasks (including an *n*-back incubation period), symmetry span, and one of the excluded insight problems (see footnote 1 regarding the excluded problem-solving tasks), in that order. In the second session, they completed a demographic survey, one of the excluded analytical-filler problems, operation span, the other excluded insight problem, the other version of the alternative uses task (including an *n*-back incubation period), and the other excluded analytical-filler problem, in that order. We counterbalanced the insight and alternative uses tasks such that half the subjects completed Coins and Brick in Session 1 and half completed Pigpen and Knife in Session 1 (the incubation *n*-back tasks were not counterbalanced, and were encountered in this order: countries, body parts, instruments, and colors).

## Scoring

### *Working memory tasks*

We scored both operation span and symmetry span using partial credit scoring (Conway et al., 2005), in which the total number of items recalled in correct serial position was summed across the task. These scores were then individually converted into *z* scores based on our database of 3,393 UNCG students, and then averaged to create a WMC composite. We retained task data from only those participants who scored  $\geq 85\%$  accuracy on the processing component of the span task, as is conventional, in order to be sure that retained subjects were not ignoring the processing portion of the span task and treating the task as a “simple” or “short-term memory” span task requiring only memory rehearsal (e.g., Conway et al., 2005; Redick et al., 2012).

### *TUT rates*

We categorized all off-task thought reports (i.e., tuning out and zoning out) as mind wandering, and calculated TUT rate separately for each task; TUT rate refers in each task to the proportion of thought probes on which a subject endorsed either tuning or zoning out.

### *Alternative uses tasks*

Three raters (the first author and two other UNCG psychology graduate students who have published research on creativity) scored each subject’s individual responses on a scale of 1–5. For scoring purposes, the raters were told to view creative ideas as having three facets: they are uncommon but apt, they are remotely linked to everyday objects and ideas, and they are often clever. The raters saw an alphabetized spreadsheet of all responses, from all subjects in Experiments 1 and 2, presented without any identifying information. Each spreadsheet of responses was alphabetized differently (i.e., A–Z, Z–A, and M–Z/A–L) to eliminate the possibility that responses later in the alphabet would be scored differently. We asked the raters to read the entire list of responses before they started rating responses in the order they received them. Finally, raters were asked to use the entire range of scores while rating, and after they were finished, the raters sorted the responses by score to ensure that this was true. Raters showed adequate agreement (Brick  $\alpha = .688$ , Knife  $\alpha = .689$ ) that was consistent with prior work using such ratings (e.g., Silvia, 2008a).

After all ratings were complete, we calculated scores for each subject in three ways: (a) an *average* score, by taking an average across all of a subject's responses; (b) a *top-two* score, by averaging across the two best responses selected by the subject; (c) a *max-two* score, by averaging across the two responses that were the most highly rated. For all of these divergent creativity scores, ratings were first averaged across raters for each response, and then across all of the responses for each person. Prior research indicates that average scoring is a bit more reliable than top-two scoring, but top-two scoring may have greater validity than average scoring, insofar as it better predicts creative personalities (Silvia, 2008a). We added max-two scoring as another means by which to assess subjects' best ideas, but without requiring optimal discernment, which might also covary with executive control capabilities.

## Results

For all analyses, we report null hypothesis significance tests with an alpha of .05 and, where appropriate, Cohen's *d* or partial eta squared ( $\eta^2$ ) to estimate effect size.

## Subjects

Data from subjects were omitted from analyses for scoring less than 85% on the processing component of either complex span task (18 people) and exceptionally poor performance on the SART or *n*-back tasks with embedded thought probes (4 people); "poor performance" outliers were determined by collapsing nontarget accuracy across *n*-back tasks and excluding anyone who had an accuracy of  $\leq 75\%$  in either the collapsed *n*-back or on the nontarget trials of the stand-alone SART (*M* accuracy rates and RTs appear in Appendix C). 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% African American, 5.8% Asian, 5.0% Multiracial, and 2.5% Other; for ethnicity, reported separately, 5% self-identified as Latino/Hispanic. Subjects completed both sessions within a mean of 15.2 (*SD* = 12.9) days.

## Primary analyses

Here we present analyses that are central to our hypotheses; we address secondary questions and exploratory analyses in the subsequent section.

## Descriptive statistics

Table 1 presents descriptive statistics for WMC and TUT rates, and Table 2 presents the correlations among these variables. Operation span and symmetry span correlated fairly well, allowing us to collapse their *z* scores into a WMC composite, which was normally distributed with a mean close to zero (indicating a sample representative of our larger database). On average, subjects reported TUTs to about 40% of task probes, whether the task stood alone (i.e., SART) or represented an incubation task (i.e., *n*-backs); a one-way analysis of variance (ANOVA) indicated no statistical differences in TUT rates across tasks,  $F(4, 592) = 0.71, p = .58, MSE = 0.06, \eta^2 = .01$ . Furthermore, TUT rates correlated significantly across all pairwise tasks,  $r_s = .24$  to  $.65$ , suggesting that mind wandering was also reasonably consistent across people. WMC, however, did not correlate consistently with TUT rate. Significant negative correlations arose from the *n*-back incubation tasks within the two insight tasks, but not from the

alternative uses tasks or the stand-alone SART (the latter of which failed to replicate prior findings from our laboratory; McVay & Kane, 2009, 2012a).

**Table 1.** *Experiment 1 Descriptive Statistics for Working Memory Capacity Tasks and Mind-Wandering Rates*

Task	N	M	SD	Min	Max	Skew	Kurtosis
Sspan Z	120	-.062	.975	-2.742	1.832	-.396	-.094
Ospan Z	120	.061	.949	-3.279	1.546	-.941	1.258
WMC Z	120	-.001	.782	-1.637	1.589	-.253	-.613
SART TUT	119	.367	.196	.033	.800	.203	-1.072
Coins TUT	120	.413	.260	0	1	.272	-.750
PigPen TUT	119	.399	.252	0	1	.396	-.489
Brick TUT	120	.397	.236	0	1	.438	-.462
Knife TUT	119	.413	.261	0	1	.432	-.575

Note. Z = z scores based on database of more than 3,000 people; WMC = working memory capacity; SART = sustained attention response task; TUT = task-unrelated thought proportion.

**Table 2.** *Experiment 1 Correlations Among Working Memory Capacity Tasks and Mind-Wandering Rates*

Variable	1	2	3	4	5	6	7	8	9
1. Sspan Z	1.00								
2. Ospan Z	.32**	1.00							
3. WMC Z	.82**	.81**	1.00						
4. SART TUTs	-.04	-.01	-.03	1.00					
5. Coins TUTs	-.22	-.15	-.23*	.40**	1.00				
6. Pigpen TUTs	-.15	-.15	-.19*	.45**	.48**	1.00			
7. Brick TUTs	-.04	-.05	-.06	.36**	.65**	.50**	1.00		
8. Knife TUTs	-.03	-.10	-.08	.48**	.41**	.61**	.24**	1.00	
9. Incubation TUTs	-.14	-.17	-.19*	.54**	.81**	.83**	.75**	.73**	1.00

In Experiment 1, we first assessed the consistency of divergent thinking scores across tasks, and found that overall creativity ratings (collapsed across both pre- and postincubation) in the Brick task correlated well with overall creativity in the Knife task, across average scoring,  $r(117) = .58, p < .001$ , top-two scoring,  $r(117) = .38, p < .001$ , and max-two scoring,  $r(117) = .59, p <$

.001. Therefore, for all subsequent alternative-uses analyses, we averaged scores across Brick and Knife tasks to create our divergent creativity measures. Note, however, that we did so separately for preincubation task periods and postincubation task periods, thus creating preincubation divergent thinking scores and postincubation divergent thinking scores (and also allowing use to calculate a pre-to-post incubation change score).

*Is executive control helpful or harmful for creativity?*

We first operationalize executive control variation via WMC scores. WMC did not correlate with postincubation divergent-thinking scores, for average,  $r(114) = .01, p = .90$ , top-two,  $r(114) = .07, p = .47$ , or max-two,  $r(114) = .03, p = .72$  scoring. To consider executive control variation as reflected by TUT propensity, we averaged the TUT rates across the two alternative-uses incubation periods. Mind wandering did not correlate with postincubation divergent creativity scores for average,  $r(114) = -.08, p = .42$ , top-two,  $r(114) = -.10, p = .29$ , or max-two,  $r(114) = .01, p = .92$ , scoring methods. To make closer contact with Baird et al. (2012), whose primary analyses were on percent improvement in divergent creativity scores from pre- to postincubation, we also computed change as they did:  $[(\text{postincubation divergent creativity} - \text{preincubation divergent creativity}) / \text{preincubation divergent creativity}] \times 100$ . Incubation TUT rate did not correlate with creativity percent change: average,  $r(114) = -.08, p = .38$ , top-two,  $r(114) = -.07, p = .45$ , max-two,  $r(114) = .02, p = .85$  (nor did TUTs correlate with a simple pre-to-post difference score, with  $r_s = -.09$  to  $.02, p_s = .33$  to  $.83$ ).

Although both WMC and TUT rate showed null associations with divergent thinking, people who are high in WMC *and* TUT rates may be more creative than others. To test whether the WMC  $\times$  TUT interaction predicted divergent thinking, we used average postincubation creativity scoring as the outcome measure in a hierarchical linear regression. WMC and alternative-uses-incubation TUT rate entered at Step 1 and the interaction of the two (WMC  $\times$  TUT) entered in at Step 2. The interaction effect was not significant (see Table 3).

**Table 3.** *Experiment 1 WMC  $\times$  TUT Hierarchical Linear Regression on Divergent Thinking*

						95% CI for B	
Predictors	B	SE	$\beta$	$t$	$p$	Lower	Upper
Block 1 ( $\Delta R^2 = .006$ )							
WMC z score	.002	.029	.006	.068	.946	-.056	.059
DT TUT Rate	-.023	.029	-.075	-.792	.430	-.080	.034
Constant	2.044	.029		71.459	.000	1.987	2.101
Block 2 ( $\Delta R^2 = .010$ )							
WMC $\times$ TUT	.034	.031	.102	1.086	.280	-.028	.095

*Note.* WMC = working memory capacity; DT = divergent thinking; TUT = task-unrelated thoughts;  $N = 116$ ; all predictors are  $z$  scores based on the final sample (for centering purposes).

## Secondary and exploratory analyses

Here we asked whether executive control variation predicted preincubation creativity. It did not: Neither WMC nor TUT rate had significant effects. Preincubation divergent thinking scores did not correlate with WMC [average scoring:  $r(117) = .14, p = .14$ ; top-two scoring:  $r(117) = .13, p = .18$ ; max-two scoring:  $r(117) = .04, p = .66$ ]. Examining TUT rates based on preincubation performance is particularly informative because it can tell us whether subjects mind-wander strategically if they are given a reason to do so (i.e., if they might productively mind-wander to help generate new alternative uses of a brick). They did not. Preincubation divergent thinking scores did not correlate with incubation-period TUT rates [average scoring:  $r(117) = -.04, p = .65$ ; top-two scoring:  $r(117) = -.10, p = .29$ ; max-two scoring:  $r(117) = -.02, p = .85$ ].

## Discussion

The main questions that Experiment 1 addressed, from an individual-differences perspective, were whether executive control is helpful or harmful for postincubation creative thinking. WMC did not correlate with postincubation creativity (or preincubation creativity, or pre-to-post change in creativity) as assessed by ratings of divergent-thinking output. Our findings for incubation-period TUT rates were also clear and consistent: Mind wandering showed no association with divergent thinking. Moreover, being higher in both WMC and mind-wandering propensity, which might allow one to regulate their off-task thinking toward productive ends, did not improve divergent-thinking performance. Finally, mind wandering did not increase for subjects who had not yet provided high-quality alternate uses in the divergent thinking tasks. In general, then, we find no evidence that executive control—or its failure—contributes to individual differences in the creativity of divergent thinking (whether before or after incubation).

We found, but did not expect, that WMC-TUT correlations were inconsistent across tasks: Whereas higher WMC predicted lower TUT rates during *n*-back incubation in the insight problems (with typically modest magnitudes), they did not within the alternative uses incubation tasks or the stand-alone SART; in the multilevel model that combined TUT rates across all 5 tasks, the WMC effect on TUTs did not reach conventional significance ( $\beta = -0.22, SE = .14, p = .11$ ). We do not know whether these discrepant findings across tasks reflect meaningful, systematic differences between contexts that elicit versus suppress a WMC-TUT association, or whether we are simply dealing with random variation around a small but true negative effect (note that none of the correlations here were positive). As noted by Kane and McVay (2012), long-duration SARTs tend to elicit small WMC-TUT correlations, in the range of  $r = -.20$  (McVay & Kane, 2009, 2012a), and shorter-duration SART-like tasks elicit still weaker correlations that only reach magnitudes of  $-.20$  at the latent-variable level, reflecting shared variance across multiple WMC and TUT-rate assessments (McVay & Kane, 2012b). Given that all the tasks here were in the short-duration range of 20 min (vs. 40+ min), we consider it most likely that WMC has a generally weak association with TUTs during tasks that require executive control to block distractions or regulate habitual responding, and that tasks that do not stress these control processes to their limits by requiring them over long periods of time will be less likely to detect it. Although it requires more data from more laboratories to effectively test such a

claim, a recent meta-analysis (Randall et al., 2014) indicated that WMC and other cognitive ability measures show a significant but small negative correlation with laboratory TUT rates ( $\rho = -.14 [-.09 - -.19]$ ) and that this correlation is significantly stronger in tasks that are longer than 30 min ( $\rho = -.20 [-.15 - -.24]$ ) versus shorter than 30 min ( $\rho = -.07 [.01 - -.15]$ ).

We designed Experiment 2 to replicate our null WMC results and to better understand our null mind-wandering results. That is, Experiment 2 explored whether our null TUT-creativity associations derived from the particular incubation task that we used. We chose the *n*-back for Experiment 1, not only because it has been used successfully as an incubation task before (Dijksterhuis & Meurs, 2006), but also because it elicits substantial TUT rates (McVay et al., 2013); indeed, subjects in Experiment 1 reported mind wandering on approximately 40% of *n*-back probes, with considerable variation around that mean. At the same time, the Sio and Ormerod (2009) meta-analysis suggested that incubation tasks with lower cognitive load were most likely to show benefits, and a 2-back task arguably presents a higher-than-ideal load (see also Baird et al., 2012). Of course, if incubation benefits require low cognitive load to occur, but mind wandering can be substantial even in high load tasks like the 2-back, then it suggests that mind wandering is not the active ingredient behind creative incubation. In any case, because the Sio and Ormerod meta-analysis also indicated that reading science fiction frequently elicits significant incubation (e.g., Smith & Blankenship, 1991), Experiment 2 incubation tasks required word-by-word reading of science fiction texts as a means to give mind wandering its best opportunity to show some creative benefits.

## **Experiment 2**

The procedure matched that in Experiment 1, but we changed the incubation task and we assessed two personality dimensions relevant to creative problem solving (e.g., Schooler & Melcher, 1995; Silvia et al., 2008): Openness to experience (from the Openness subscale of the NEO-PI-3; McCrae & Costa, 2010) and Need for Cognition (from the Need for Cognition questionnaire; Cacioppo, Petty, & Kao, 1984). These measures addressed novel hypotheses about interest and motivation. People who score high on personality scales of Openness tend to be intellectually curious and are motivated to be creative (McCrae, 1987), and those who score high on Need for Cognition tend to choose to engage in critical thinking and enjoy it when they do (Cacioppo & Petty, 1982). We thus used these measures to indicate intrinsic motivation and valuing creative goals—two factors that could positively impact performance (for a meta-analysis, see de Jesus, Rus, Lens, & Imaginário, 2013). Specifically, in separate moderator analyses we tested whether the impact of WMC, mindwandering, or both, on divergent creativity would be stronger for subjects who were higher in Openness or Need for Cognition.

Additionally, recall that Baird et al. (2012) found that a retrospective measure of daydreaming modestly correlated with divergent thinking. Although we had begun data collection when that study was published, we were able to take advantage of the “fantasy” facet within the Openness to experience measure. The fantasy facet is conceptually similar to what was measured in the Baird et al. daydreaming measure, and so we investigated the correlations among our fantasy measure, in-the-moment TUT reports, and divergent-thinking measures.



## **Method**

### **Subjects**

To match our sample sizes across experiments, while accounting for likely data loss, we aimed to collect complete datasets from 120–130 people in both sessions. Once again, introductory psychology students at UNCG earned partial credit toward a course requirement for participating in each of two 120 min sessions (earning more credit for Session 2 than Session 1). Again, using two complete semesters as our data-collection stopping rule, 154 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.

### **Working memory span tasks**

Working memory tasks and assessments were identical to those in Experiment 1.

### **Mind-wandering assessment: Ongoing tasks and thought probes**

#### *Reading tasks*

For each incubation task and the stand-alone task for TUT assessment, subjects read a different science fiction story excerpt, presented one word at a time on-screen, and responded by pressing the space bar whenever they detected an anomaly. Specifically, for the reading tasks, we used the first 1200 words from Chapter 1 of the following novels: *Across the Universe* (Beth Revis), *Divergent* (Veronica Roth), *Rex Rising* (Chrystalla Thoma), *Whatever Became of the Squishies?* (Claire Chilton), and *True Hero?* (Jack Hessey). We also used the first 67 words of *The House of Scorpion* (Nancy Farmer) as practice trials for the stand-alone reading task.

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 read normally but to respond to any anomaly only after they had seen the second swapped word (e.g., following “*with*”). Nearly identical to the *n*-back procedure in Experiment 1, each word appeared for 400 ms, followed by a 600-ms blank screen (we had also noted that Smallwood, McSpadden, and Schooler [2008] reported  $M = 304$  ms per word in a word-by-word reading study). We presented words at a constant pace, rather than allowing subjects to read at their leisure, to control for individual differences in reading speed and to ensure 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, for 35 probes total (Experiment 1 SART had presented 60; Experiment 1 *n*-backs each presented 21). After each story, subjects answered five multiple-choice questions, via mouse click, to motivate their reading for comprehension.

#### *Mind-wandering probes*

Thought probes and their instructions were identical to those in Experiment 1.

### **Divergent creativity: Alternative uses tasks**

Divergent thinking tasks and assessments were identical to those in Experiment 1.

## Noncognitive assessments

We combined the Openness and Need for Cognition scales into one computer-presented questionnaire, 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; the sequence then repeated until all 66 items were included. All responses used a 5-point Likert-type scale (i.e., labeled with *strongly disagree*, *disagree*, *neutral*, *agree*, and *strongly agree*).

### *Openness*

This questionnaire, taken from the NEO-PI-3 (McCrae & Costa, 2010), comprises six facets: (a) fantasy (e.g., “I enjoy concentrating on a fantasy or daydream and exploring all its possibilities, letting it grow and develop”); (b) aesthetics (e.g., “I am intrigued by the patterns I find in art and nature”); (c) feelings (e.g., “I experience a wide range of emotions or feelings”); (d) actions (e.g., “I think it’s interesting to learn and develop new hobbies”); (e) ideas (e.g., “I often enjoy playing with theories or abstract ideas”), and (f) values (e.g., “I consider myself broad-minded and tolerant of other people’s lifestyles”). Eight items defined each facet, with half reverse-coded. Although we were primarily interested in the general openness construct, we were secondarily interested in the fantasy facet, which, at face value, measures engagement in fantastical thinking and daydreaming.

### *Need for cognition*

We 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 (e.g., “I find satisfaction in deliberating hard and for long hours”; “I prefer my life to be filled with puzzles that I must solve”; “Learning new ways to think doesn’t excite me very much” [reverse scored]). Although the original questionnaire used a -4 to +4 Likert-type scale, we used a -2 to +2 scale to keep the responses consistent with the openness measure in the combined questionnaire (i.e., *strongly disagree*, *disagree*, *neutral*, *agree*, and *strongly agree*).

## General procedure

As in Experiment 1, we tested subjects individually across two 2-hr sessions with the same computers. In Session 1, subjects completed the stand-alone reading task, one of the alternative uses tasks, symmetry span, and one of the excluded insight problems (see Footnote 1 regarding excluded problem-solving measures). In Session 2, they completed a demographic survey, an excluded analytic problem, operation span, another excluded insight problem, the other alternative uses task, and finally, the noncognitive questionnaire containing Openness and Need for Cognition scales. Subjects completed the tasks in the aforementioned order for each session.

## Results

### Subjects

Data from 15 subjects were omitted for failing the processing-portion criterion (85% accuracy on the processing component) on either span task and from one subject who was older than our

target age range of 18 to 30 years. Using the same incubation-task-accuracy criterion as in Experiment 1, outliers were determined by collapsing nontarget accuracy across all reading tasks (both stand-alone and incubation), but here, no one's accuracy was 75% or less and so we retained data from all remaining subjects (Appendix D presents mean accuracy and RTs). Data from 115 subjects were included for analyses, who ranged in age from 18–28 years ( $M = 18.84$ , 67.0% female). By self-report, the final sample's racial composition was 55.7% White, 31.3% African American, 3.5% Asian, 5.2% Multiracial, 1.7% Native American or Alaskan Native, and 2.6% Other; for ethnicity, reported separately, 4.3% self-identified as Latino/Hispanic. Subjects completed both sessions within a mean of 18.9 ( $SD = 17.4$ ) days.

### Primary analyses

As in Experiment 1, we begin with analyses that are central to our hypotheses; the subsequent section addresses secondary questions and exploratory analyses.

### Descriptive statistics

Tables 4 and 5 present descriptive statistics for WMC and TUT measures, and for the noncognitive measures, respectively. Table 6 presents correlations among the WMC and TUT measures and Table 7 presents the correlations between the noncognitive measures and the WMC and TUT variables. As in Experiment 1, we collapsed across the two WMC tasks (operation span and symmetry span) to create one WMC  $z$  score composite measure; these WMC  $z$  scores were lower in Experiment 2 ( $M = -0.19$ ,  $SD = 0.83$ ) than in Experiment 1 ( $M = 0.00$ ,  $SD = 0.78$ ),  $t(233) = 1.82$ ,  $p = .07$ , 95% CI  $[-0.02, 0.40]$ ,  $d = 0.24$ , but this near-significant difference reflected only a small effect.

**Table 4.** *Experiment 2 Descriptive Statistics for Working Memory Capacity Tasks and Mind-Wandering Rates*

Task	N	M	SD	Min	Max	Skew	Kurtosis
Sspan Z	115	-.311	1.035	-2.615	1.959	-.262	-.641
Ospan Z	115	-.071	1.004	-3.145	1.546	-.844	.573
WMC Z	115	-.191	.826	-.2783	1.585	-.507	.124
Stand-alone TUT	115	.257	.213	0	.861	.937	.213
Coins TUT	114	.265	.228	0	1	.950	.446
PigPen TUT	114	.279	.252	0	.944	1.128	.401
Brick TUT	115	.298	.245	0	1	.931	.170
Knife TUT	115	.317	.263	0	1	.719	-.268

*Note.* Z =  $z$  scores based on database of over 3,000 people; WMC = working memory capacity; TUT = task-unrelated thought proportion.

**Table 5.** *Experiment 2 Descriptive Statistics for Noncognitive Variables*

Variable	N	M	SD	Min	Max	Skew	Kurtosis
Fantasy Facet	115	.615	1.109	-1.543	3.022	-.050	-.721

Aesthetics Facet	115	.755	.996	-2.000	2.561	-.320	-.396
Feelings Facet	115	.602	1.067	-2.049	2.829	-.465	.185
Actions Facet	115	.467	1.169	-2.194	3.917	.399	.201
Ideas Facet	115	.428	.845	-1.259	2.630	.294	.130
Values Facet	115	.349	1.077	-2.643	2.833	-.073	-.003
Openness	115	.810	1.050	-2.242	3.134	.003	.114
Need for Cognition	115	-.046	.993	-2.729	2.836	.132	-.011

*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 6.** *Experiment 2 Correlations Among Working Memory Capacity Tasks and Mind-Wandering Rates*

Variable	1	2	3	4	5	6	7	8	9
1. Span Z	1.00								
2. Ospan Z	.31**	1.00							
3. WMC Z	.82**	.80**	1.00						
4. Stand-alone TUTs	-.03	.18	.09	1.00					
5. Coins TUTs	-.06	-.05	-.07	.53**	1.00				
6. Pigpen TUTs	.01	.18	.11	.56**	.64**	1.00			
7. Brick TUTs	.03	.06	.05	.56**	.71**	.58**	1.00		
8. Knife TUTs	.01	.12	.08	.60**	.48**	.72**	.40**	1.00	
9. Incubation TUTs	-.01	.10	.06	.67**	.84**	.89**	.80**	.80**	1.00

*Note.* Z = *z* scores based on database of more than 3,000 people; WMC = working memory capacity; TUT = task-unrelated thought proportion; Incubation tasks were collapsed across all four reading-incubation tasks. *N* = 113–115. \*\**p* < .01.

**Table 7.** *Experiment 2 Correlations of Noncognitive Variables With Working Memory Capacity Tasks and Mind-Wandering Rates*

Variable	Fantasy Facet	Aesthetics Facet	Feelings Facet	Actions Facet	Ideas Facet	Values Facet	Openness	Need for Cognition
1. Span Z	.05	.05	-.13	-.22*	.04	-.04	-.05	-.15
2. Ospan Z	-.06	-.17	-.13	-.23*	-.16	-.13	-.21*	-.13
3. WMC Z	.00	-.07	-.16	-.28**	-.07	-.10	-.16	-.17
4. Stand-alone TUTs	.14	-.04	.03	-.03	-.03	.04	.03	.00
5. Coins TUTs	.13	.10	.07	.20*	-.02	.02	.12	.00
6. Pigpen TUTs	.02	.07	-.05	.13	-.12	.03	.02	-.09
7. Brick TUTs	-.02	.01	.05	-.04	-.19*	-.07	-.06	-.10
8. Knife TUTs	.15	.13	-.01	.17	-.09	.18	.13	-.12

9. Incubation TUTs	.07	.08	.01	.14	-.13	.04	.05	-.11
10. Average Pre Inc DT	.28**	.22*	.18	.29**	.06	.21*	.30**	.03
11. Average Post Inc DT	.24*	.12	.17	.22*	.09	.22*	.26**	.01
12. Top-Two Pre Inc DT	.18	.10	.15	.13	-.04	.05	.14	-.04
13. Top-Two Post Inc DT	.17	.01	.17	.26**	.09	.14	.20*	.01
14. Max-Two Pre Inc DT	.25*	.17	.16	.18	-.02	.02	.19*	.00
15. Max-Two Post Inc DT	.21*	-.02	.09	.13	.03	.04	.11	-.03

*Note.*  $Z = z$  scores based on database of more than 3,000 people; WMC = working memory capacity; TO = tune outs; ZO = zone outs; TUT = task-unrelated thoughts; DT = divergent thinking; Pre Inc = pre-incubation; Post Inc = post-incubation.  $N = 108-115$ . \*  $p < .05$ . \*\*  $p < .01$ .

On average, across all reading tasks, subjects reported TUTs to approximately 30% of thought probes, whether the reading task stood alone or as an incubation task. Just as in Experiment 1, a one-way ANOVA did not indicate TUT rate differences across tasks,  $F(4, 568) = 1.18, p = .32, MSE = 0.06$ . Furthermore, TUT rates again correlated substantially across all pairwise tasks,  $r_s = .40$  to  $.72$ , indicating between-subjects stability. Overall, TUT rates were significantly lower here, in the reading tasks, than in the Experiment 1 SART and  $n$ -back tasks, and this remained true whether we took the average TUT rate across all five tasks,  $t(229) = 4.52, p < .001, 95\% CI [0.06, 0.16], d = 0.59$ , or compared the tasks one-by-one: Stand-alone,  $t(232) = 4.12, p < .001, 95\% CI [0.06, 0.16], d = 0.54$ ; Brick,  $t(233) = 3.15, p = .002, 95\% CI [0.04, 0.16], d = 0.41$ ; Knife:  $t(232) = 2.80, p = .006, 95\% CI [0.03, 0.16], d = 0.37$ . Negative correlations between WMC and TUT rate were of inconsistent magnitude and statistical significance across tasks in Experiment 1, but here, in apparent contrast, WMC tended toward weak and nonsignificant positive correlations with TUT rate.

Overall divergent creativity scores (collapsed across pre- and postincubation) in the Brick task correlated well with those in the Knife task, regardless of using average,  $r(111) = .56, p < .001$ , top-two,  $r(111) = .40, p < .001$ , or max-two scoring,  $r(111) = .50, p < .001$ . As in Experiment 1, then, we averaged Brick and Knife scores from preincubation periods to create a single preincubation divergent creativity measure for each subject, and the Brick and Knife scores from postincubation periods to create a single postincubation divergent creativity measure for each subject (for each scoring system).

#### *Is executive control helpful or harmful for creativity?*

We first consider WMC as an indicator of executive control. WMC failed again to significantly predict postincubation alternative uses scores, for average  $r(106) = .04, p = .72$ ; top-two,  $r(106) = -.03, p = .76$ ; max-two  $r(106) = .04, p = .72$ .

We next consider mind-wandering propensity as a measure of executive control. Also consistent with Experiment 1, TUT rates during alternative uses incubation did not predict divergent

creativity scores, either for average,  $r(106) = .04, p = .70$ , top-two,  $r(106) = -.07, p = .45$ , or max-two scoring methods,  $r(106) = .03, p = .77$ . We also examined pre-to-post percent change in divergent creativity scores (as in Baird et al., 2012), but these did not correlate with incubation TUT rates, either: average  $r(104) = .02, p = .83$ ; top-two  $r(104) = -.02, p = .82$ ; max-two  $r(104) = -.05, p = .63$  (nor did TUT rates correlate with a simple pre-to-post difference score, with  $r_s = -.11$  to  $-.02, p_s = .26$  to  $.81$ ).

Once again, we tested via hierarchical linear regression whether the interaction between WMC and TUT rates could predict divergent creativity, with the possibility that people high in WMC and TUT rates may be more creative. Table 8 illustrates that it did not.

**Table 8.** Experiment 2 WMC  $\times$  TUT Hierarchical Linear Regression on Divergent Thinking

Predictors	B	SE	$\beta$	t	p	95% CI for B	
						Lower	Upper
Block 1 ( $\Delta R^2 = .002$ )							
WMC z score	.011	.034	.032	.331	.741	-.056	.078
DT TUT Rate	.012	.033	.034	.352	.725	-.054	.077
Constant	2.033	.033		61.136	.000	1.967	2.099
Block 2 ( $\Delta R^2 = .006$ )							
WMC $\times$ TUT	-.028	.035	-.079	-.795	.428	-.097	.042

*Note.* WMC = working memory capacity; DT = divergent thinking; TUT = task-unrelated thoughts.  $N = 108$ . All predictors are  $z$  scores based on the final sample (for centering purposes).

*Does intellectual motivation moderate the association between WMC and creativity?*

Before testing our hypothesis, we first examined the reliabilities of the questionnaires and the correlations among the measures. Internal consistency scores were good: Coefficient alpha for the Need for Cognition scale (16 items) was .81 and for the Openness scale (48 items) was .89. Openness scores correlated positively with Need for Cognition scores,  $r(113) = .42, p < .001$  (see also Sadowski & Cogburn, 1997), but nonsignificantly in the opposite direction with WMC,  $r(113) = -.16, p = .10$ ; Need for Cognition's negative correlation with WMC was of similarly modest magnitude,  $r(113) = -.17, p = .07$  (see also Benny & Banks, 2015; Hill et al., 2013). Regarding divergent creativity, Openness significantly predicted average score,  $r(106) = .26, p < .01$  (see Figure 2), and top-two score,  $r(106) = .20, p < .05$ , but not the max-two score,  $r(106) = .11, p = .25$ . In contrast, Need for Cognition did not correlate significantly with any measure of postincubation divergent creativity [average  $r(106) = .01, p = .94$ ; top-two,  $r(106) = .01, p = .95$ ; max-two,  $r(106) = -.03, p = .80$ ]. We also analyzed pre-to-post percent change in divergent creativity scores, but these did not correlate with either Openness ( $r_s = -.08$  to  $.08, p_s = .42$  to  $.59$ ) or Need for Cognition ( $r_s = -.04$  to  $.05, p_s = .62$  to  $.81$ ). Thus, although subjects who were more open to experience tended to be more divergently creative than were those who were less open, their divergent creativity did not benefit any more from an

incubation period and its attendant mind wandering.

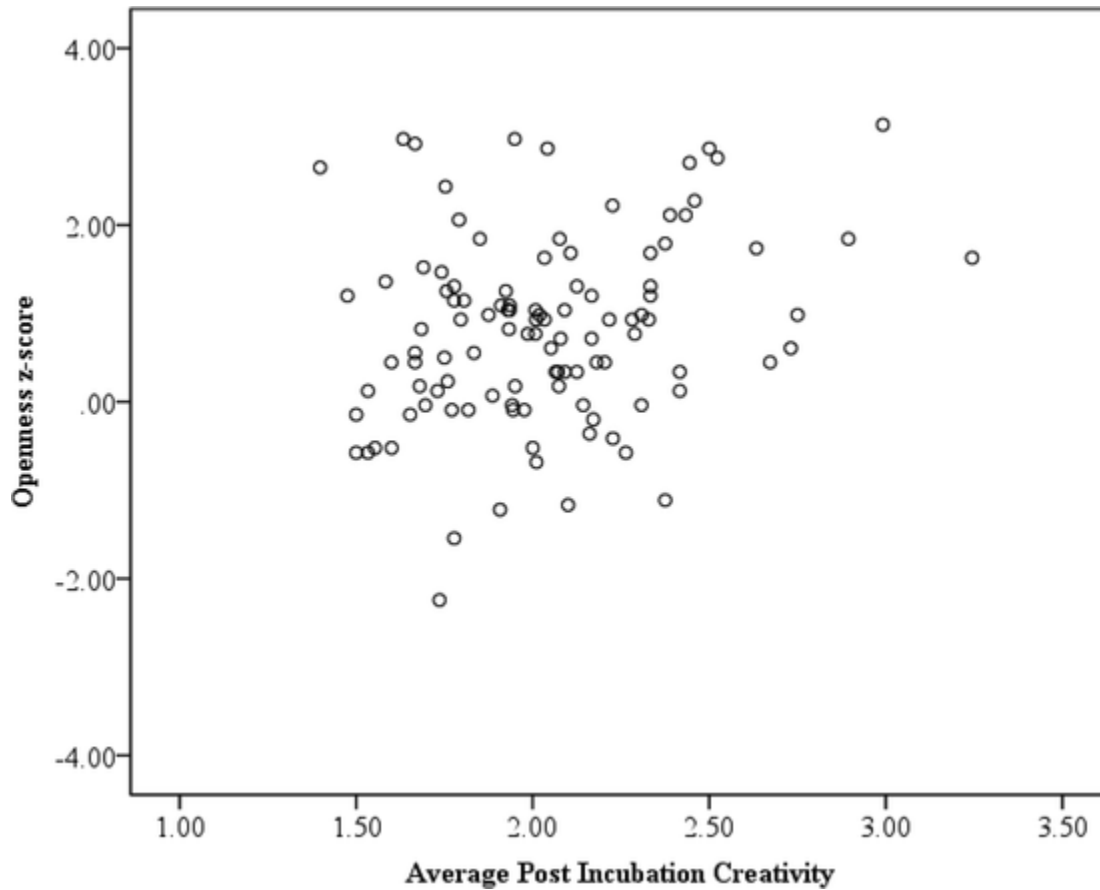


Figure 2.

Experiment 2 correlation scatterplot between openness and divergent creativity (average scoring).

Several regressions tested our hypothesis that Openness and Need for Cognition might interact with WMC to predict divergent thinking. In our first hierarchical linear regression, we entered WMC and Openness in Block 1 to predict average postincubation divergent creativity, and added their interaction at Block 2 (see Table 9). Likewise, in our second hierarchical linear regression, we entered WMC and Need for Cognition in Block 1 and their interaction in Block 2 (see Table 10). Neither Openness nor Need for Cognition moderated the relation between WMC and average postincubation divergent creativity.

Table 9. Experiment 2 WMC × Openness Hierarchical Linear Regression on Divergent Thinking

Predictors	B	SE	$\beta$	t	p	95% CI for B	
						Lower	Upper
Block 1 ( $\Delta R^2 = .072$ )							
WMC z score	.026	.033	.076	.797	.427	-.039	.091
Openness	.092	.032	.269	2.832	.006	.027	.156
Constant	2.033	.032	-	63.402	.000	1.970	2.097

Block 2 ( $\Delta R^2 = .000$ )							
WMC × Openness	.002	.031	.006	.063	.950	-.059	.063

*Note.* WMC = working memory capacity.  $N = 108$ . All predictors are  $z$  scores based on the final sample (for centering purposes).

**Table 10.** *Experiment 2 WMC × Need for Cognition Hierarchical Linear Regression on Divergent Thinking*

Predictors	B	SE	$\beta$	t	p	95% CI for B	
						Lower	Upper
Block 1 ( $\Delta R^2 = .001$ )							
WMC $z$ score	.013	.034	.038	.380	.705	-.055	.081
Need for Cognition	.005	.034	.014	.139	.889	-.063	.072
Constant	2.033	.033		61.114	.000	1.967	2.099
Block 2 ( $\Delta R^2 = .010$ )							
WMC × Need for Cog	-.033	.033	-.100	-1.003	.318	-.098	.032

*Note.* WMC = working memory capacity; Need for Cog = Need for Cognition.  $N = 108$ . All predictors are  $z$  scores based on the final sample (for centering purposes).

### Secondary and exploratory analyses

#### *Did executive control variation predict preincubation creative cognition?*

As in Experiment 1, neither WMC nor TUT rate showed any effects here. Preincubation divergent creativity scores did not correlate with WMC [average scoring:  $r(109) = -.04, p = .69$ ; top-two scoring:  $r(109) = .08, p = .38$ ; max-two scoring:  $r(109) = .03, p = .75$ ] or with TUT rates during the divergent-thinking incubation tasks [average scoring:  $r(109) = .07, p = .47$ ; top-two scoring:  $r(109) = .00, p = .97$ ; max-two scoring:  $r(109) = .13, p = .17$ ].

#### *Is the fantasy facet of Openness associated with incubation TUTs or creative cognition?*

As noted earlier, the Openness subscale comprises six facets (McCrae & Costa, 2010), and we were especially interested in the fantasy-TUT correlation. 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), the fantasy facet had face validity in measuring a propensity to mind-wander in daily life, or, at least, an imaginative engagement with such experiences. In fact, although the fantasy facet (8 items) showed adequate reliability ( $\alpha = .75$ ), it was not correlated with probed TUT rates collapsed across all incubation reading tasks,  $r(111) = .07, p = .44$ .

Regarding the creative aspect of the fantasy scale, and consistent with the daydreaming questionnaire findings from Baird et al. (2012), fantasy-facet scores correlated modestly but significantly with postincubation divergent creativity based on average and max scoring [ $r(106)$



= .24,  $p = .01$ , and  $r(106) = .21$ ,  $p = .03$ , respectively] and trended similarly for top-two scoring,  $r(106) = .17$ ,  $p = .07$ . These fantasy correlations are of similar magnitude to those from the overall Openness score, however, which suggests that it is not daydreaming propensity, *per se*, that drives the association between Openness and divergent creativity.

## Discussion

We designed Experiment 2 as a conceptual replication of Experiment 1 that also addressed new questions about creative cognition and personality. Our main findings were that, as in Experiment 1, neither WMC nor incubation mind wandering predicted creative responses in divergent thinking tasks. Thus, despite modifying our incubation tasks to present science-fiction texts, in order to reduce their cognitive demands and align them with “best practices” (Sio & Ormerod, 2009), we were still unable to demonstrate any beneficial effects of TUTs on divergent thinking. Indeed, as in Experiment 1, any nonsignificant trends suggested by the data tended toward the *negative*, with higher TUT rates weakly predicting lower creative performance.

The null associations between WMC and TUT rate in Experiment 2 bear consideration, particularly because the nonsignificant WMC-TUT correlations trended positive in Experiment 2, rather than negative as in Experiment 1. Both McVay and Kane (2012b) and Unsworth and McMillan (2013) found that higher WMC subjects reported fewer TUTs than did lower WMC subjects while reading expository or fictional texts in more traditional formats (i.e., not presented word by word in a context of anomaly detection). The current null-to-positive WMC-TUT correlations conflict with those prior reports. As in Experiment 1, however, we cannot be certain whether a lack of significant (and negative) WMC effects indicates a systematic boundary condition around the WMC-TUT association, such as the relatively short durations of these tasks, or simply random variation around a true-but-weak negative association between these constructs in nature (Randall et al., 2014).

But what should we make of the null correlation between WMC and divergent thinking, now seen across two experiments? On one hand, this finding arguably follows from the perspective that a lack of cognitive control might benefit creativity (e.g., Aiello et al., 2012; Chrysikou & Thompson-Schill, 2011), insofar as WMC correlates positively with so many other domains of higher-order cognitive ability, but not this one. On the other hand, a null WMC-creativity association is surprising because divergent thinking correlates positively with fluid intelligence (e.g., Silvia, 2008a, 2008b; Nusbaum & Silvia, 2011), which shares at least half its variance with WMC (e.g., Kane et al., 2005; Oberauer et al., 2005). The strongest evidence for an effect of fluid intelligence on divergent thinking comes from latent-variable models, where the variance common to the intelligence measures, and the variance common to the divergent creativity measures, are statistically isolated from the method or rater variance that also contribute to the individual task scores. So, to provide the most rigorous test for any WMC-creativity association, we collapsed data across both experiments and used structural equation modeling to create two models of divergent creativity—one using average scoring, and the other using top two scoring (see Figure 3). In both models, WMC was indicated by operation span and symmetry span, with the two paths constrained to be equal and the WMC variance set to 1. Likewise, Divergent Thinking was indicated by Brick and Knife and those two paths were set to be equal with the

Divergent Thinking variance set to 1. Finally, Brick and Knife were each indicated by the three raters' scores.

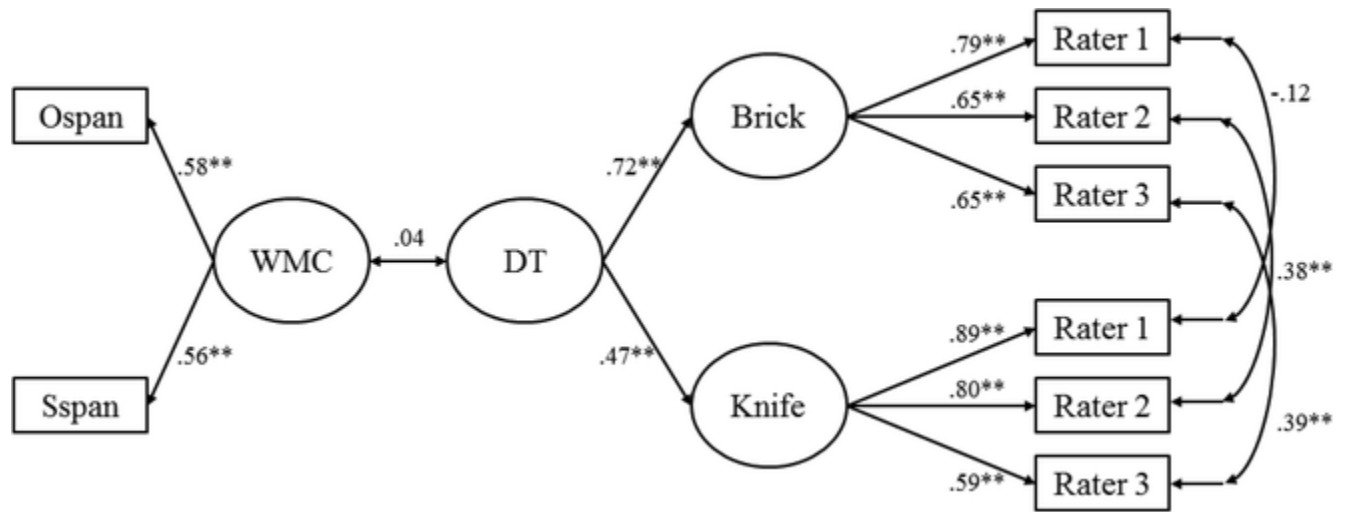


Figure 3. Structural equation model with data collapsed across Experiments 1 and 2 (average scoring). Ospan = operation span; Sspan = symmetry span; WMC = working memory capacity; DT = divergent thinking. \*\*  $p < .01$ .

Both models of WMC and divergent thinking had good fit, regardless of using average scoring,  $\chi^2(16) = 26.392, p = .049$ ; CFI = 0.978; RMSEA = 0.053, 90% CI = [0.004, 0.087]; SRMR = 0.045, or top two scoring,  $\chi^2(16) = 24.257, p = .084$ ; CFI = 0.980; RMSEA = 0.047, 90% CI = [0.00, 0.083]; SRMR = 0.039. In both models, all paths were statistically significant except for our correlation of interest—the relation between WMC and Divergent Thinking (average scoring:  $\beta = .04, p = .79$ ; top two scoring:  $\beta = .06, p = .62$ ). These modeling findings provide compelling evidence that WMC does not influence divergent creativity. As noted above, although a view that executive control may be harmful to creativity should predict a negative correlation between WMC and divergent thinking, the null correlations we observe suggest, at least, that executive control is not helpful and this finding stands in stark contrast to others on higher-order cognitive functions. We suggest that future research should explore the surprising discrepancy indicated here between the predictive powers of WMC and fluid intelligence, by having subjects complete tests of all three constructs and testing whether the fluid intelligence variance that is not shared with WMC is that which predicts variation in divergent thinking.

Experiment 2 explored new hypotheses about personality, primarily whether intellectual motivation—as measured by Openness to Experience and Need for Cognition—moderated the relation between WMC and creative cognition (see Dollinger, 2003; McCrae, 1987). Although we found that Openness predicted postincubation divergent thinking, neither it nor Need for Cognition acted as a moderator. Thus, WMC's null associations with divergent thinking in Experiments 1 and 2 did not likely arise from a lack of curiosity or motivation in our student sample.

Our exploratory assessments of the Openness fantasy facet indicated that it did not correlate with TUT rates during ongoing tasks, but it did modestly predict divergent thinking. We suggest that

the fantasy measure did not correlate with TUTs, in part, because fantasy-facet items ask more about one's feelings toward, and active approaches to, daydreaming than about its sheer frequency; fantasy scores may thus better reflect one's imaginative engagement with fantasy and daydreams than one's propensity to engage in mind wandering while performing other tasks. If so, it follows that people who are more creative generally will also be more creative with their own subjective experiences. Note that this plausible interpretation turns the presumed directionality of the creativity–mind wandering association on its head—rather than mind wandering boosting creativity, perhaps creativity boosts imaginative or constructive mind wandering. As well, the fact that the overall Openness score (which comprises not only fantasy-proneness but also intellectual curiosity, aesthetic appreciation, emotionality, and open-mindedness) predicted divergent thinking as well as did the fantasy-specific facet score, suggests that daydreaming propensity or quality is not central—and thus not causal—in the association between Openness and creativity. Among its other goals described below, Experiment 3 tested whether two more straightforward retrospective measures of mind-wandering frequency would correlate more strongly with probed TUT rates, and whether they would also predict variation in creative cognition (as in Baird et al., 2012).

### **Experiment 3**

As mentioned previously, while data collection for our Experiment 2 was ongoing, Baird et al. (2012) published results suggesting that greater mind wandering is correlated with greater divergent creativity. Baird et al. compared different types of incubation activities—a demanding task, an undemanding task, unstructured rest, or no break—during divergent thinking. They found that only the undemanding incubation-task group improved in creativity scores from pre- to postincubation. Importantly, this same group reported the most mind wandering during incubation, according to a retrospective questionnaire that immediately followed the incubation period. Thus, Baird et al. presented *correlational* evidence that two different dependent measures, mind wandering and divergent-thinking improvement, were both increased by the experimental manipulation of reducing incubation-task demand. Divergent thinking scores also correlated positively with scores on a general measure of daydreaming propensity in daily life (the Daydreaming Frequency subscale of the Imaginal Process Inventory; Singer & Antrobus, 1972). These experimental and individual-differences findings led Baird and colleagues to conclude that mind wandering facilitates creative problem solving, although it's important to understand that their design did not actually provide experimental evidence for a causal effect of mind wandering on creativity.

Our first two experiments might be considered failed “conceptual replications” of the Baird et al. (2012) study, insofar as they provided no individual-differences support for the claim that mind-wandering enables divergent creativity. Indeed, we take Underwood's (1975) perspective that individual-differences findings should serve as early falsification tests for emerging theory, and so our consistent null results provide a significant refutation of the notion that creative cognition (at least as indicated by divergent thinking) arises from uncontrolled, off-task thought (see also Singer & Antrobus, 1963). We must acknowledge, however, that we may have failed to find evidence of TUT-inspired creativity because our incubation tasks were akin to the Baird et al.

“demanding” task, which elicited no apparent mind-wandering benefit (our incubation tasks did, however, produce substantial mind-wandering rates, with *M*s of .30–.40 and rates of .50–.60 within +1 *SD*, and so if mind wandering actually drives incubation benefits, we should have seen them). The Baird et al. demanding task was a 12 min, numerical version of the 1-back task, in which rare colored digits prompted a 1-back parity judgment (black digits required no response); our 20 min 2-back tasks from Experiment 1 were almost certainly more demanding than this. Although it’s less clear that our reading-anomaly tasks from Experiment 2 were as challenging as the Baird et al. “demanding” 1-back task, they do seem more demanding than their “undemanding” task of responding to rare colored digits with a concurrent parity judgment (a 0-back task).

In Experiment 3, then, we replaced the 20 min incubation tasks with the 12 min undemanding digit task from Baird et al. (2012), who promptly provided their materials to us. We also replaced the noncognitive measures from Experiment 2 with the Daydreaming Frequency Scale (as in Baird et al.) and the Mind Wandering Scale from the same instrument (which not only measures frequency of mind-wandering experiences, but also frames some questions around succeeding/failing to concentrate on ongoing tasks). These questionnaires allowed us to test whether general retrospective measures of mind-wandering frequency and activity predict divergent thinking (as in Baird et al., and arguably consistent with Experiment 2’s fantasy-scale findings) and whether they predicted in-the-moment TUT rates (which fantasy scores did not predict).

Despite our using the Baird et al. (2012) undemanding incubation task and our adjusting the timing of task events to mimic theirs, Experiment 3 was not intended as a “close” replication of that study because we focused on correlational rather than experimental effects, and we thought it important to maintain critical aspects of our prior design. The most significant differences between our studies were the following: (a) we kept our online thought probes to measure in-the-moment TUTs during the incubation period (rather than relying on Baird’s single retrospective questionnaire following incubation, which is vulnerable to memory biases and mental aggregation errors); (b) we assessed divergent thinking in the alternative uses task using subjective ratings (rather than using the uniqueness of each response as Baird’s criterion for creativity, which is a problematic dependent measure; see Silvia et al., 2008); (c) we explicitly instructed subjects to be creative in their responding rather than to simply generate many responses, which is critical to actually assessing creative thinking rather than memory fluency (see Nusbaum et al., 2014); (d) we had subjects complete only one alternative uses task rather than several; and (e) we included measures of WMC and a Mind Wandering scale.

## **Method**

### **Subjects**

Just as we did in our previous experiments, we aimed to collect data from approximately 130 subjects (here, across 2 semesters worth of time that spanned half of a Fall, a full Spring, and half a Summer semester). In exchange for participating in a single 60 min session, 132 UNCG undergraduates earned either partial credit toward an introductory psychology course

requirement or \$15 (five of these 132 participants did not complete the entire session). Paid participants ( $N = 40$ ) were recruited by posting flyers in classroom buildings around campus, particularly in those that housed fine arts, music, and theater departments and courses, in an effort to recruit more prototypically “creative” students than might otherwise enroll in introductory psychology.

### **Working memory span tasks**

We used the same two complex span tasks as before, operation span and symmetry span, to measure working memory capacity.

### **Mind-wandering assessment: Ongoing tasks and thought probes**

We measured mind wandering using the same probes as in the previous two experiments. Here, however, subjects completed no stand-alone task and only one incubation task (for the one alternative uses task), which took about 12 min to complete.

#### *Undemanding 0-back task*

A series of numbers (1–8) appeared on screen, one at a time, for 1000 ms, followed by a fixation cross for 1,500 ms. If the number appeared in green, participants responded to this target trial by pressing the *F* key if the number was odd or the *J* key if the number was even (4% of the 250 trials). If the number appeared in black, subjects withheld responding (96% of trials).

#### *Mind-wandering probes*

Subjects responded to 14 thought probes appearing unpredictably during the 0-back incubation task. This thought-probe rate was comparable to those in the Experiments 1 and 2 incubation tasks, adjusting for task length. Although probes were identical to those in Experiments 1 and 2, they did not appear only after rare target stimuli.

### **Mind-wandering assessment: Retrospective questionnaires**

Subjects completed two subscales from the Imaginal Process Inventory (IPI), Daydreaming Frequency (e.g., “When I am not paying close attention to my job, a book, or TV, I tend to be daydreaming”; “I lose myself in active daydreaming”; “Whenever I have time on my hands I daydream”) and Mind Wandering (e.g., “At times it is hard for me to keep my mind from wandering”; “During a lecture or speech, my mind often wanders”; “No matter how hard I try to concentrate, thoughts unrelated to my work always creep in”). Each subscale presented 12 items, and subjects responded to each item by selecting one of five multiple-choice options. None of the Daydreaming Frequency items were reverse-scored but six of the Mind Wandering items were. Items were scored on a 1–5 scale and then summed within the subscale, with higher scores indicating more daydreaming or mind wandering, and a maximum score of 60 on each.

### **Divergent thinking task**

We used only one version of the alternative uses task, the brick task, to measure divergent creativity. In line with the time limits from Baird et al. (2012), subjects had only 4 min to come up with as many creative uses for a brick as they could (rather than 10 min, as in Experiments 1

and 2). After the first 2 min of working on the brick task, subjects switched to the 0-back incubation task for approximately 12 min, and then returned to the brick task for another 2 min. Similar to Experiments 1 and 2, at the end of the session, subjects selected their top two answers among their preincubation responses, postincubation responses, and all of their responses. Because we only used the brick version of the divergent thinking task, and the procedure of that task also differed from the previous two experiments, raters scored these responses separately from the responses in Experiments 1 and 2. The raters and rating procedure, however, were the same as in the previous experiments.

## **General procedure**

Subjects completed the 60 min session in groups of 1 to 4. The same computers as in Experiments 1 and 2 presented all tasks. Subjects first completed a demographic questionnaire, followed by operation span, the alternative uses task (brick), symmetry span, and finally, the IPI questionnaire subscales.

## **Results**

### **Subjects**

Data from 20 subjects were eliminated for processing-portion errors on either WMC task (i.e., processing accuracy below 85%, as in Experiments 1 and 2). Using the same criterion as in Experiments 1 and 2, outliers were determined by looking at incubation-period nontarget accuracy, but similar to Experiment 2, no one had accuracy  $\leq 75\%$ , and so all data from remaining subjects were retained (mean 0-back accuracy and RTs can be found in Appendix E). Data from 107 subjects were included for analyses, who ranged in age from 18–29 years ( $M = 19.63$ , 61.7% female). By self-report, the final sample's racial composition was 52.3% White, 29.0% African American, 7.5% Asian, 7.5% Multiracial, and 2.8% Other. Additionally, regarding self-reported ethnicity (reported separately), 7.5% identified as Latino/Hispanic. Of the 40 people recruited from flyers around campus, 33 had viable data and, of those, eight reported seeing the flyer in a “creative” building (i.e., fine arts, music, or theater).

### **Primary analyses**

As in Experiments 1 and 2, we begin with hypothesis-driven confirmatory analyses; the subsequent section reports secondary and exploratory analyses.

### **Descriptive statistics**

Table 11 presents descriptive statistics for WMC, TUT, and questionnaire measures; Table 12 reports their correlations. As before, we averaged operation span and symmetry span to create a WMC  $z$  score composite measure. Mean WMC  $z$  scores in this sample of for-credit and for-pay subjects were significantly higher in Experiment 3 ( $M = 0.24$ ,  $SD = 0.71$ ) than in both Experiment 1 ( $M = -0.00$ ,  $SD = 0.78$ ),  $t(225) = 2.37$ ,  $p = .02$ , 95% CI  $[-0.43, -0.04]$ ,  $d = 0.32$ , and Experiment 2 ( $M = -0.21$ ,  $SD = 0.84$ ),  $t(221) = 4.24$ ,  $p < .001$ , 95% CI  $[-0.65, -0.24]$ ,  $d = 0.57$ .

Table 11. *Experiment 3 Descriptive Statistics for Working Memory Capacity Tasks, Mind-Wandering Rates, and Questionnaire Measures*

Task	N	Mean	SD	Min	Max	Skew	Kurtosis
Sspan Z	107	.228	.969	-1.853	1.959	-.318	-.844
Ospan Z	107	.243	.763	-2.140	1.546	-.697	-.111
WMC Z	107	.235	.708	-1.742	1.518	-.492	-.468
Brick TUT	107	.475	.250	0	1	-.142	-.735
Daydreaming Q	105	40.524	8.916	21	59	-.227	-.434
Mind Wandering Q	105	38.657	6.251	24	56	.387	.400

Note. Z = z scores based on database of more than 3000 people; WMC = working memory capacity; TUT = task-unrelated thoughts; Q = questionnaire.

Table 12. *Experiment 3 Correlations Among Working Memory Capacity Tasks, Mind-Wandering Rates, and Questionnaire Measures*

Variable	1	2	3	4	5	6
1. Sspan Z	1.00					
2. Ospan Z	.33*	1.00				
3. WMC Z	.86**	.76**	1.00			
4. Brick TUTs	.10	-.02	.06	1.00		
5. Daydreaming Q	.00	.16	.09	.21*	1.00	
6. Mind Wandering Q	-.08	-.02	-.06	.29*	.51**	1.00

Note. Z = z scores based on database of more than 3,000 people; WMC = working memory capacity; TUT = task-unrelated thought proportion; Q = questionnaire.  $N = 105-107$ . \* $p < .05$ .

\*\*  $p < .01$ .

TUT rates during the Brick task were higher in the 0-back task in Experiment 3 ( $M = 0.47$ ,  $SD = 0.25$ ) than in the 2-back task in Experiment 1 ( $M = 0.40$ ,  $SD = 0.24$ ),  $t(225) = 2.41$ ,  $p = .02$ , 95% CI  $[-0.14, -0.01]$ ,  $d = 0.32$ , and the reading-anomaly task in Experiment 2 ( $M = 0.30$ ,  $SD = 0.25$ ),  $t(221) = 5.39$ ,  $p < .001$ , 95% CI  $[-0.24, -0.11]$ ,  $d = 0.72$ . As in Experiment 2, WMC correlations with TUT rates varied nonsignificantly around 0, but trended toward weakly positive. Unlike the Openness or Fantasy facet scores from Experiment 2, both the Daydreaming and Mind Wandering subscales correlated positively and significantly with probed TUT rates. Finally, interrater reliability was adequate for the brick divergent thinking task, with a coefficient alpha of .721.

*Is executive control helpful or harmful for creativity?*

Once again, WMC did not predict postincubation divergent creativity, regardless of whether we used average,  $r(102) = -.03$ ,  $p = .79$ , top-two,  $r(102) = .00$ ,  $p = .96$ , or max-two scoring,  $r(102) = .03$ ,  $p = .79$ . Also replicating our prior findings, mind wandering during incubation did not measurably benefit postincubation divergent creativity, either for average scoring,  $r(102) = -.05$ ,  $p = .63$ , top-two scoring,  $r(102) = .02$ ,  $p = .82$ , or max-two scoring,  $r(102) = -.08$ ,  $p = .40$ ; moreover, just as in Experiments 1 and 2, percent change in divergent creativity from pre- to

postincubation was not correlated with incubation TUT rates: average  $r(101) = -.09, p = .39$ , top-two  $r(101) = .05, p = .61$ , max-two  $r(101) = -.09, p = .38$  (nor did TUT rates correlate with a simple pre-to-post difference score, with  $r_s = -.09$  to  $.05, p_s = .38$  to  $.61$ ).

Finally, we tested whether the Mind Wandering or Daydreaming Frequency subscales predicted postincubation divergent creativity. All Mind Wandering subscale correlations were nonsignificant and very near zero: average  $r(100) = .01, p = .91$ ; top-two  $r(100) = -.02, p = .81$ ; max-two  $r(100) = -.06, p = .57$ . The Daydreaming subscale, in contrast, showed weak positive correlations with divergent thinking, but none of them quite reached conventional levels of significance: average  $r(100) = .14, p = .17$ ; top-two  $r(100) = .20, p = .05$ ; max-two  $r(100) = .14, p = .16$ . Moreover, Daydreaming scores did not correlate significantly with pre-to-post incubation changes in creativity: average  $r(100) = -.03, p = .80$ ; top-two  $r(100) = .10, p = .31$ ; max-two  $r(100) = .03, p = .79$ , suggesting that daydreaming did not influence creative incubation.

Just as we had done previously, we again tested whether the interaction between WMC and mind wandering predicted divergent creativity, that is, whether those high in both WMC and TUT rate might be most creative. We tested this via three separate regressions, using incubation TUT rate, Daydreaming subscale score, and Mind Wandering subscale score (see Table 13). For all three, we first entered WMC and a mind-wandering measure at Step 1, and then added the interaction between the two at Step 2. None of the mind-wandering indicators interacted with WMC to predict average postincubation divergent creativity.

**Table 13.** *Experiment 3 Hierarchical Linear Regressions on Divergent Thinking*

Predictors	B	SE	$\beta$	t	p	95% CI for B	
						Lower	Upper
WMC $\times$ Brick TUT hierarchical linear regression (N = 104)							
Block 1 ( $\Delta R^2 = .003$ )							
WMC z score	-.010	.043	-.023	-.234	.815	-.095	.075
Brick TUT Rate	-.020	.042	-.047	-4.68	.641	-.103	.064
Constant	2.249	.042	-	53.703	.000	2.166	2.332
Block 2 ( $\Delta R^2 = .001$ )							
WMC $\times$ TUT	-.011	.041	-.028	-.275	.784	-.093	.070
WMC $\times$ Daydreaming subscale hierarchical linear regression (N = 102)							
Block 1 ( $\Delta R^2 = .022$ )							
WMC z score	-.024	.040	-.059	-.592	.555	-.104	.056
Daydreaming score	.058	.040	.144	1.443	.150	-.022	.137
Constant	2.240	.040	-	56.601	.000	2.162	2.319
Block 2 ( $\Delta R^2 = .014$ )							
WMC $\times$ Daydreaming	-.053	.044	-.123	-1.208	.230	-.140	.034
WMC $\times$ Mind Wandering subscale hierarchical linear regression (N = 102)							
Block 1 ( $\Delta R^2 = .002$ )							
WMC z score	-.017	.041	-.041	-.412	.681	-.097	.064



Mind Wandering score	.004	.040	.010	.103	.918	-.076	.084
Constant	2.240	.040		56.015	.000	2.161	2.320
Block 2 ( $\Delta R^2 = .014$ )							
WMC $\times$ Mind Wandering	-.048	.041	-.119	-1.160	.249	-.130	.034

*Note.* WMC = working memory capacity; TUT = task-unrelated thoughts during brick incubation; Daydreaming = score on the daydreaming subscale; Mind wandering = score on the Mind Wandering subscale. All predictors are  $z$  scores based on the final sample (for centering purposes).

### Secondary and exploratory analyses

As in Experiments 1 and 2, we asked whether executive control variation predicts preincubation creative cognition. The answer was again, “no.” Preincubation divergent creativity ratings were not predicted by WMC [average:  $r(104) = .02, p = .85$ ; top-two:  $r(104) = .08, p = .43$ ; max-two:  $r(104) = .05, p = .64$ ] or incubation-period TUT rate (average:  $r(104) = .09, p = .37$ ; top-two:  $r(104) = -.04, p = .72$ ; max-two:  $r(104) = .05, p = .63$ ).

### Discussion

Experiment 3 modified the Baird et al. (2012) incubation methods to bring our procedures more in line with one that appeared to demonstrate the benefits of mind wandering for creative cognition. Specifically, we used their low-demand (0-back) incubation task—which was nearly half the length of our incubation periods from Experiments 1 and 2—in addition to shortening our alternative uses tasks and including the Daydreaming Frequency subscale of the IPI. At least with respect to individual differences in divergent creativity, we once again failed to find any significant facilitation from incubation-period mind wandering: The rate at which subjects reported TUTs during the incubation task did not correlate with their postincubation creative output (or with their pre-to-post change in creative output) in the divergent thinking task. Our most important finding, then, is that people who mind wander more during an undemanding incubation task produce no more creative output in a divergent thinking task than do those who mind wander less. This individual-differences finding (which replicates those from Experiments 1 and 2), provides compelling evidence against theoretical claims that mind wandering enhances creative thought.

The only positive evidence from Experiment 3 for the ostensible benefits of mind wandering was a near-significant correlation ( $p = .05$ ) between top-two-scored divergent thinking and the IPI Daydreaming subscale ( $r = .20$  [.01, .38]), which was of similar magnitude to that from Baird et al. (2012;  $r = .22$  [.06, .37]). Although the other two scoring methods did not yield significant correlations with Daydreaming ( $r_s = .14$  [-.06, .33]), we consider the top-two result to represent a positive conceptual replication of Baird et al. because it very closely matched the original effect size despite our sample being a bit smaller (see Simonsohn, 2015). At the same time, Daydreaming scores did not correlate with the pre-to-post changes in creativity with incubation (see also Baird et al.), and so the questionnaire findings seem to disconfirm the theoretical claim that incubation benefits derive from mind wandering.

The Daydreaming results suggest that people who report more off-task thinking in everyday life are also a bit more creative overall, but the similar IPI Mind Wandering subscale did not correlate with any of our divergent thinking indices ( $r_s = -.06$  to  $.01$ ). So, mind-wandering propensity, itself, does not seem linked to creative thought. We therefore interpret the questionnaire data, along with the fantasy-facet findings from Experiment 2, to indicate that divergent creativity is not associated with simply more off-task thinking, but rather with a certain kind of off-task thinking—namely, one that is mainly positive, intentional and, perhaps also, creative (McMillan et al., 2013; Singer, 1966). Indeed, several of the Daydreaming items that were most strongly correlated with divergent thinking scores reflected a propensity for off-task thought when otherwise idle: “Whenever I have time on my hands I daydream” ( $r = .18$ ); “When I am at a meeting or show that is not very interesting, I daydream rather than pay attention” ( $r = .16$ ); “On a long bus, train, or airplane ride I daydream” ( $r = .23$ ).

In contrast, divergent thinking was uncorrelated with a Daydreaming item about off-task thought at work or school (“I daydream at work (or school)”);  $r = .02$ , or with several work- or school-related items from the Mind Wandering scale: “My mind seldom wanders while I am working” ( $r = .03$ ); “My mind seldom wanders from my work” ( $r = -.03$ ); “During a lecture or speech, my mind often wanders” ( $r = -.12$ ); “No matter how hard I concentrate, thoughts unrelated to my work always creep in” ( $r = .01$ ); “During a speech, meeting, or lecture, I often ‘come to,’ realizing that I have not heard a word the speaker was saying” ( $r = .00$ ). Creative people do not struggle, then, to maintain on-task thoughts while engaged in demanding activities; rather, creative people may choose to engage in daydreaming when it will not otherwise be costly. From this perspective, however, daydreaming does not necessarily provide fodder for creative cognition, but rather a creative talent or outlook may allow or encourage more positive-constructive daydreaming, particularly during cognitive “down time”. The causal direction of any modest link between creativity and retrospective reports of daily life mind wandering, then, is currently ambiguous—if there is, in fact, any causal link between them. A third variable like openness to experience may cause them both, given that we found general Openness scores, which comprise much more than just fantasy and daydreaming, to predict divergent thinking as well as did the fantasy-specific facet.

Regarding individual differences in WMC, our findings indicated no association with postincubation divergent thinking quality (or with pre-to-post increases in divergent thinking), and no interaction of WMC with mind-wandering in predicting divergent creativity, replicating Experiments 1 and 2. Our data also showed no WMC correlation with incubation-task TUT rate, which replicates prior findings that, during simple vigilance tasks requiring little executive control, WMC is unrelated to mind-wandering vulnerability (Baird, Smallwood, & Schooler, 2011; McVay & Kane, 2012a).

## **General Discussion**

Our three individual-differences studies provide little evidence for claims that executive-control processes are helpful for, or harmful to, creative cognition; in short, indices of executive-control variation were uncorrelated with performance on divergent-thinking tasks. We specifically tested the potential influence of normal variation in WMC and mind-wandering propensity on creative

problem solving that was interrupted—and to be facilitated—by a brief incubation period. Of most importance, WMC did not correlate with postincubation divergent thinking scores (in all three experiments); we also found (in all three experiments) that rates of self-reported TUTs during incubation tasks were not correlated with quality of divergent thinking, despite using incubation tasks of varying demand (and yielding varying TUT rates) across experiments. The null effects we produced in predicting divergent creativity are credible and meaningful because: (a) our samples were adequately powered to detect moderate effect sizes; (b) the measures we used to assess working memory capacity, mind wandering, and divergent thinking were valid, being motivated by theory and representing prototypical instruments in their respective research literatures; moreover, our measures were reliable, showing substantial correlations with other tasks of the same construct (e.g., between WMC tasks, across incubation-period TUT assessments, between divergent thinking tasks); (c) we observed the same patterns of null results across three separate experiments; (d) the null correlations were set within a nomological net of significant correlations, such as probed TUT rate with Daydreaming and Mind Wandering questionnaires (Experiment 3), and divergent-thinking scores with Openness and fantasy scales (Experiment 2) and with a Daydreaming questionnaire (Experiment 3).

The general retrospective questionnaires about daydreaming and fantasy experiences (and, more generally, Openness to experience) provided the only significant results related to the potential benefits of mind wandering for creative cognition. In Experiment 2, the fantasy facet of Openness did not correlate with probed TUT rates during incubation, but it did positively predict postincubation divergent thinking ratings. In Experiment 3, questionnaire measures of Daydreaming and Mind Wandering were both correlated with incubation-period TUT rates, but only the Daydreaming scale—which, like the fantasy measure—likely reflected *quality* of off-task thinking rather than just quantity—modestly predicted postincubation divergent creativity (consistent with the Daydreaming questionnaire findings reported by Baird et al., 2012). In summary, then, neither WMC nor immediate-probed TUT reports, our two best indicators of executive-control variation, told us anything about our subjects' divergent thinking capabilities (again, despite the WMC tasks, TUT measures, and creativity assessments being reliable and well correlated within-constructs). Note also that none of these retrospective measures correlated with the *increase* in creative quality observed between pre- and postincubation (nor did the Daydreaming scale in Baird et al.), and so they do not reflect any kind of benefit of in-the-moment, off-task thinking, for facilitating creative thought.

Our questionnaire measures of fantasy and off-task thought suggested that better divergent thinkers also claim to engage in more fantasy-driven and intentional mind wandering in their everyday lives. Of importance, however, our data cannot tell us whether the latter indicates that certain conscious experiences are more likely to be the causes, or the effects, of one's creative potential. Given that Openness/fantasy correlated with postincubation divergent creativity without correlating with incubation TUT rate, and given that the Daydreaming, but not the Mind Wandering, questionnaire correlated with postincubation divergent creativity, and given that none of these questionnaires correlated with pre-to-post incubation increases in divergent creativity, we suggest that it is more likely that being creative causes people to have more elaborate and playful daydreams than it is that having more off-task or stimulus-independent

thoughts causes people to be more creative or insightful. Indeed, because our findings showed that better divergent thinkers were also more *generally* open to experience than were poorer divergent thinkers, even beyond fantasy and daydreaming facets, other aspects of openness are likely to be the causal mechanisms behind any association between creativity and endorsement of fantasy and daydreaming.

### **WMC and Creative Cognition**

Across all three experiments, and in the cross-experimental analyses of Experiments 1 and 2, we found that WMC did not predict postincubation divergent creativity (or preincubation creativity, for that matter). Previously published studies have reported mixed results regarding the correlation between WMC and divergent thinking, with some showing a positive association (Benedek et al., 2014; De Dreu et al., 2012) and others showing null results (Lee & Therriault, 2013; Lin & Lien, 2013, although the latter's studies were likely underpowered). It is not yet clear what accounts for these discrepancies. Our findings (and others') of null correlations between WMC and divergent thinking are surprising, however, given the established relations between fluid intelligence and divergent thinking (Nusbaum & Silvia, 2011; Silvia, 2008a) and fluid intelligence and WMC (e.g., Kane et al., 2005; Oberauer et al., 2005). If fluid intelligence predicts creativity based on the variance that it does not share with WMC, it may suggest nonexecutive contributions to the intelligence-creativity link (but see Gilhooly et al., 2007; Nusbaum & Silvia, 2011).

An anonymous reviewer suggested that the long duration of our divergent thinking tasks may have been important. Beaty and Silvia (2012) asked subjects to generate creative uses for a brick for 10 min, and found not only that the responses became more creative over time, but also that Gf interacted with time in predicting creativity scores. Specifically, people with higher Gf scores showed stable (and high) creativity over time, but people with lower Gf scores showed steeper positive slopes in creativity (starting low but increasing with time). Perhaps by providing 10 min for divergent thinking in Experiments 1 and 2 (5-min preincubation and 5-min postincubation), we allowed our lower WMC subjects to "catch up" to our higher WMC subjects. We see a few challenges for this claim. First, De Dreu et al. (2012) found robust associations between WMC and several divergent thinking indices in an even longer, 16 min brainstorming task. Second, our Experiment 3 allowed subjects only 4 min for divergent thinking (2 min preincubation and 2 min postincubation) and WMC did not predict creativity scores there. Third, in all three of our experiments, WMC failed to correlate with divergent thinking during the initial preincubation periods, which were 5 min, 5 min, and 2 min long in Experiments 1 to 3, respectively. Our findings do not appear to be driven, then, by task length. Future work should follow up by testing the differential predictive power of WMC and fluid intelligence for creative cognition, across a range of different tasks for each construct, and perhaps including convergent creativity measures, as well.

### **Mind Wandering and Creative Cognition**

We assessed the association between mind-wandering propensity and divergent thinking for two reasons. First, normal variation in TUTs during ongoing tasks is often associated with executive

control, with higher mind-wandering rates associated with poorer control (e.g., McVay & Kane, 2009; Unsworth & McMillan, 2013), allowing us to further test the control-creativity link. Second, divergent thinking improves following incubation periods that are characterized by modest cognitive demands (Sio & Ormerod, 2009), and thus may allow mind wandering (Baird et al., 2012), which, as a relatively unconstrained and sometimes “playful” form of thought, may allow mental access to more remote ideas as fodder for creative solutions (for variations on this view, see Baars, 2010; Freud, 1908; Immordino-Yang, Christodoulou, & Singh, 2012; McMillan et al., 2013; Schooler et al., 2011; Smallwood & Schooler, 2006). Both sets of findings lead to the prediction that people with higher incubation TUT rates should also show more creative thinking. All three of our experiments failed to support that critical theoretical prediction (Underwood, 1975), finding no correlation between incubation TUTs and divergent thinking quality. People who mind-wandered frequently generated no more (or less) creative uses of bricks and knives than did those who mind-wandered infrequently.

Of course, our critical TUT-rate results conflict with the findings reported by Baird et al. (2012), who found that subjects in their least demanding incubation task condition (compared to three other incubation groups), retrospectively reported the most incubation mind wandering and also showed the most improvement in divergent creativity scores from pre-to-post incubation. In effect, Baird et al. demonstrated a correlation between two previously well-known findings—that easier tasks promote higher TUT rates, and easier tasks promote greater incubation benefit—within a single study design. This already highly cited study (187 Google Scholar citations as of September 18, 2015) seems to support the claim that mind wandering facilitates creative, divergent thinking, but as we noted earlier, their study was not designed to provide experimental evidence for their causal claim that mind wandering affects creativity; instead, they found that their manipulated variable (incubation task demand) had correlated effects on their two dependent variables, retrospective mind-wandering reports and divergent creativity-score increase.

We did not design our individual-differences studies to experimentally manipulate incubation-task demands to evaluate corresponding changes in TUTs or creativity, but our Experiments 1 and 2 did present a kind of “natural experiment” by using different incubation tasks in the context of an otherwise identical design. Recall that Experiment 1 presented a 2-back task during incubation and Experiment 2 presented an anomaly detection task within word-by-word reading of science fiction stories. Using a set of *t* tests to compare the alternative uses tasks and the incubation periods therein, we found that TUT rates were significantly higher in Experiment 1 ( $M = .41$ ,  $SD = .20$ ) than in Experiment 2 ( $M = .31$ ,  $SD = .21$ ),  $t(233) = 3.74$ ,  $p < .001$ , 95% CI [0.05, 0.15],  $d = 0.49$ ; this percent increase in TUTs from Experiment 2 to Experiment 1 is considerably larger than that found between the Baird et al. demanding and undemanding incubation tasks. However, contrary to the Baird et al. (2012) results, pre- to postincubation divergent creativity change did not differ between our experiments and, if anything, trended in the wrong direction (Experiment 1:  $M = +10.27$ ,  $SD = 14.09$ ; Experiment 2:  $M = +9.15$ ,  $SD = 16.07$ );  $t(221) = 0.55$ ,  $p = .58$ , 95% CI [-2.86, 5.10],  $d = 0.07$ ).

Why did our incubation-task findings differ from those of Baird et al. (2012)? Although both studies inserted incubation tasks into a divergent thinking task, there are key differences between them. We contend that all of these differences make our methods superior for testing the theoretical claims of interest:

1. We measured mind wandering during the incubation tasks, themselves, where off-task thought has been proposed to have its causal effects, and we used in-the-moment, online thought probes that minimize memory biases. Baird et al., in contrast, used a single retrospective measure, asking subjects *after* the incubation period to estimate how frequently their minds had wandered during the incubation task (on a 1–5 scale).
2. Our study actually tested the critical correlation between incubation TUT rates and creative output, and we found it to be zero; we assume that Baird et al. did not report the comparable correlations from their study because they were underpowered within each of their experimental conditions.
3. Although the divergent thinking tasks were very similar across studies, the instructions and, thus, what we likely measured in subjects, differed. Our instructions emphasized to subjects that the creative *quality* of their output was most important, whereas Baird et al. appeared to have emphasized that the *quantity* of their creative output was most important; we (and others) argue that responses based on creativity instructions better capture the creative-thinking construct than do those based on fluency (Harrington, 1975; Nusbaum et al., 2014).
4. The two studies scored divergent thinking differently, and thus likely differed in the constructs being measured. Our scoring was based on judgments of three blinded raters, who scored each individual response based on uncommonness, remoteness to commonplace ideas, and cleverness; Baird et al. scored responses based on uniqueness, whereby any responses that appeared only one or two times in the sample received a point. Uniqueness scoring has several serious problems (Silvia et al., 2008), only one of which was addressed by Baird and colleagues. First, uniqueness scoring is confounded with fluency, as subjects who produce more responses have more opportunity to have produced more unique responses (uniqueness and fluency correlate near .90; e.g., Torrance, 2008); to address this critique, Baird et al. also scored subjects' output for fluency and concluded that, because there were no differences in *postincubation* fluency scores among experimental groups, fluency did not drive the corresponding group differences in uniqueness (i.e., in creativity). We suggest that the Baird et al. fluency analysis does not solve all of the problems with uniqueness scoring, nor does it actually address whether their key dependent variable—*percent change* in uniqueness—was independent of *percent change* in fluency. Uniqueness scoring is also problematic because it does not take appropriateness of responses into consideration. Although uniqueness may be a necessary feature of creativity, it is not sufficient—creative solutions to problems are not only original, but also apt; random and nonsensical responses to a creative prompt may be unique, but they are not valued as creative. A final deficiency in uniqueness scoring is that it depends upon sample size, effectively

punishing more powerful designs—any particular response is less likely to be considered unique as sample size grows, perversely making it more difficult to measure creativity in the large sample sizes required by individual-differences analyses.

Our primary conclusion is therefore that mind wandering during incubation periods does not contribute to creative, divergent thinking, or individual differences therein, at least as they are measured in the laboratory. Our findings also call into question whether mind wandering is more generally beneficial to creativity. On one hand, it is logically possible that incubation periods that allow for more mind wandering actually boost everyone's creative thinking to about the same extent, suggesting that the Baird et al. (2012) conclusions are valid for cross-context comparisons, but not for individual-differences comparisons. On the other hand, although this is logically possible, it is not theoretically plausible or coherent: If the cognitive processes engaged by mind wandering actually facilitate the creative quality of one's ongoing and subsequent thoughts (as in an incubation-type context), then people who engage in more of that creativity-facilitating mind wandering should show measurable benefits in creativity relative to those who engage in less mind wandering. Given that these straightforward individual-differences predictions were contradicted by the results of our three experiments, it suggests that something else about contexts that increase mind wandering is ultimately responsible for boosting divergent thinking. That is, it seems that mind-wandering propensity is *not* causal in facilitating creative thought, and that some third variable in Baird et al. produced the relative increases in both mind wandering and divergent thinking in the undemanding-incubation condition.

We do acknowledge, in any case, that no one study, or set of studies, is conclusive. It remains possible that variants on our laboratory methods might ultimately produce positive evidence for creative facilitation by mind wandering. For example, parametrically varying the length of time between thought probes (and so providing variable time to maintain a particular train of thought), might allow for the detection of a "sweet spot" for a mind-wandering benefit; or, perhaps, experimentally manipulating subjects' motivation, or the relevance of some particular expertise, might show that some minimal emotional investment might be necessary for mind wandering to facilitate divergent thinking.

Moreover, mind wandering and creativity "in the wild" may be more closely associated than our laboratory-TUT findings indicate. Both our retrospective questionnaire data (from Openness and Daydreaming scales), and those from Baird et al., indicate that people who report experiencing more fantastical and positive-constructive daydreaming in their everyday lives perform somewhat more creatively on laboratory tasks (but see Singer & Antrobus, 1963). As we noted above, however, these associations are on the weak side (arising in only one of our three divergent-thinking dependent variables), and they also leave ambiguous whether a rich fantasy life promotes creativity or creativity promotes a rich fantasy life (or, whether a third variable like openness to experience promotes both fantasy and creativity). Finally, although creative responses to divergent thinking probes are moderately correlated with "real world" creativity (e.g., Beaty, Smeekens, Silvia, Hodges, & Kane, 2013; Silvia et al., 2008), and so laboratory creativity measures can tell us something about true creative accomplishments, it may be that the fruits of mind wandering are substantially greater when people work on creative projects that are

very important to them, and when they do so over extended periods of time that cannot be approximated by a lab setting. (The latter was loosely suggested by Schooler, Gable, Hopper, & Mrazek, unpublished; cited in Smallwood & Schooler, 2015, whose daily life study found that creative insights during off-task thinking felt more like “Aha!” experiences to professional writers and physicists, but these insights were actually of similar quality to those that came during “work” related thought and activity.)

### **WMC and Mind Wandering**

Negative correlations between WMC and TUT rates were weak in Experiment 1, evident in only some of the demanding incubation tasks. Across all of the incubation tasks in Experiments 2 and 3, WMC did not predict probe-caught mind wandering in either direction. Although the null findings from Experiment 3 were consistent with prior findings of null WMC-TUT associations during simple vigilance tasks that require little executive control (Baird et al., 2011; McVay & Kane, 2012a), our findings seem inconsistent with the general claim that superior executive abilities lead to less frequent mind wandering during ongoing tasks (e.g., McVay & Kane, 2010). As we have noted throughout, such findings suggest either some important moderators of the WMC-TUT association (such as the executive-control demands, or the length, of the ongoing task; Baird et al., 2011; Levinson et al., 2012; McVay & Kane, 2012a), or random variation around a relatively weak effect, or both. A recent meta-analytic approach to the question has supported the broad theoretical claim that WMC and other indicators of cognitive ability are negatively but weakly correlated with mind-wandering propensity during ongoing tasks, particularly if those tasks are of long enough duration to challenge sustained attention efforts (Randall et al., 2014).

### **Conclusions**

Our multi-experiment study of creative, divergent thinking following breaks for incubation yielded consistent findings: Across all three experiments, WMC showed no association to creative responding during divergent thinking tasks. We also found no evidence that TUT rates, as another index of executive-control (in)ability, have any bearing on individual differences in creative responding to divergent thinking prompts. In all three experiments, subjects who reported high rates of mind wandering when probed during incubation tasks, regardless of the particular incubation task used, were no more likely to generate creative ideas than were those who reported low rates of mind wandering. The only connection we found between off-task thought and creativity was that people who reported on retrospective questionnaires that they engaged in frequent and rich fantasy and daydreaming in their daily lives, especially during what otherwise might be considered “down time,” showed more creative divergent thinking in the lab than did those who reported rich fantasy and daydreaming. This weak association, with  $r_s \approx .20$ , seems just as likely to indicate that creativity facilitates daydreaming (or that a third variable like openness to experience facilitates both daydreaming and creativity) as it is to indicate that daydreaming facilitates creativity.

### **Footnotes**



1 In short, we found that WMC positively predicted postincubation problem solving of Pigpen in Experiment 1, but not in Experiment 2, and WMC had no association to Coins solving in either experiment; TUT rates did not predict solving either problem in either experiment. For archival purposes, we report the methodological details about these insight tasks (along with two noninsight tasks that we included to disguise these “trick” problems) and our analyses of their data, at the Open Science Framework site: <https://osf.io/4du7j/>.

2 As a secondary concern, we investigated whether the modest negative correlation between WMC and TUTs in the laboratory (often,  $r \approx -.20$ ; Kane & McVay, 2012) might vary with subjects’ efforts to concentrate (as it appears to in daily life; Kane et al., 2007). We wondered whether WMC might more strongly predict laboratory TUTs on occasions when subjects reported actually trying to focus; perhaps the literature has underestimated the correlation between WMC and mind-wandering rate because higher WMC subjects weren’t universally attempting to control thought during long and boring laboratory tasks. To test this possibility, we followed each thought probe with a separate assessment of concentration attempt (i.e., 1–7 Likert-scale rating, via keypress, for the prompt, “I was trying to concentrate on the current task”). In each experiment, we collapsed thought reports across all tasks and measured mind wandering on each probe 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), we used hierarchical linear modeling for binary outcomes with a Bernoulli distribution (Raudenbush & Bryk, 2002). We group centered the Level 1, within-subject, variable (concentration) and grand-mean centered the Level 2, between-subjects, variable (WMC). Our models included fixed effects for concentration at Level 1 and fixed effects for WMC on the intercept and on the concentration slope (i.e., a WMC  $\times$  concentration interaction) at Level 2; we also added a random effect for the intercept. Results for each experiment are presented in Appendix A. To summarize, concentration ratings negatively predicted TUTs but, contrary to our prediction, concentration did not moderate the association between WMC and TUTs. As well, a significant person-level random effect on the intercept suggests that there were individual differences in TUT rates beyond that accounted for by WMC and concentration. Our findings thus indicate that self-reported concentration did not moderate the WMC-TUT associations in the lab. Although we found substantial individual and contextual variation in concentration ratings, it may be that this variation is restricted compared to that elicited by the full range of everyday activities that young adults complete outside the lab.

3 Some of the divergent thinking analyses in Experiment 1 did not include data from all 120 people because not everyone generated responses to both tasks.

4 Some of the divergent thinking analyses in Experiment 2 did not include data from all 115 people because not everyone generated responses to both tasks.

5 Some of the divergent thinking analyses in Experiment 3 did not include all 107 people because not everyone generated responses to both tasks; in addition, for questionnaire analyses, two people did not complete the measures because they ran out of time.

## References

- Ahmed, L., & de Fockert, J. W. (2012). Focusing on attention: The effects of working memory capacity and load on selective attention. *PLoS ONE*, *7*, e43101.
- Aiello, D. A., Jarosz, A. F., Cushen, P. J., & Wiley, J. (2012). Firing the executive: When an analytic approach to problem solving helps and hurts. *The Journal of Problem Solving*, *4*, 72–83.
- Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: Experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills*, *23*, 399–417.
- Baars, B. J. (2010). Spontaneous repetitive thoughts can be adaptive: Postscript on “mind wandering.” *Psychological Bulletin*, *136*, 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*, 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*, 1604–1611.
- Batey, M., Furnham, A., & Safiullina, X. (2010). Intelligence, general knowledge and personality as predictors of creativity. *Learning and Individual Differences*, *20*, 532–535.
- 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*, 1–46.
- Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts*, *6*, 309–319.
- Beaty, R. E., Smeekens, B. A., Silvia, P. J., Hodges, D. A., & Kane, M. J. (2013). A first look at the role of domain-general cognitive and creative abilities in jazz improvisation. *Psychomusicology: Music, Mind, and Brain*, *23*, 262–268.
- Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, *46*, 73–83.
- Benny, B., & Banks, J. B. (2015). Under pressure: An examination of the predictors of choking. *Journal of Individual Differences*, *36*, 93–100.
- Bowden, E. M., & Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, *10*, 730–737.
- Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M.

J.Kane, A.Miyake, & J. N.Towse (Eds.), *Variation in working memory* (pp. 76–106). New York, NY: Oxford University Press.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 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, NY: Cambridge University Press.

Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition*, *35*, 1445–1455.

Chrysikou, E. G., & Thompson-Schill, S. L. (2011). Dissociable brain states linked to common and creative object use. *Human Brain Mapping*, *32*, 665–675.

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*, 331–335.

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*, 769–786.

Cowan, N., Elliot, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42–100.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*, 422–433.

De Dreu, C. K. W., Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working memory benefits creative insight, musical improvisation, and original ideation through maintained task-focused attention. *Personality and Social Psychology Bulletin*, *38*, 656–669.

de Jesus, S. N., Rus, C. L., Lens, W., & Imaginário, S. (2013). Intrinsic motivation and creativity related to product: A meta-analysis of the studies published between 1990–2010. *Creativity Research Journal*, *25*, 80–84.

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.

- Engel de Abreu, P. M. J., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology, 104*, 974–986.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *The Psychology of Learning and Motivation, 44*, 145–199.
- Freud, S. (1908). Creative writers and daydreaming. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 9, pp. 143–153). London, UK: Hogarth Press.
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1–16.
- Gilhooly, K. J., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking & Reasoning, 15*, 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.
- Glausiusz, J. (2011). Living in a dream world. *Scientific American Mind, March/April, 2*, 4–31.
- 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, UK: Cambridge University Press.
- Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology, 24*, 1149–1167.
- Harrington, D. M. (1975). Effects of explicit instructions to “be creative” on the psychological meaning of divergent thinking test scores. *Journal of Personality, 43*, 434–454.
- Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York, NY: Oxford University Press.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General, 136*, 217–240.
- Hill, B. D., Foster, J., Elliott, E. M., Shelton, J. T., McCain, J., & Drew Gouvier, W. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality, 47*, 22–25.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107–129.

- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 851–860.
- Immordino-Yang, M. H., Christodoulou, J. A., & Singh, V. (2012). Rest is not idleness: Implications of the brain's default mode for human development and education. *Perspectives on Psychological Science*, *7*, 352–364.
- Jarosz, A. F., Colflesh, G. H., & Wiley, J. (2012). Uncorking the muse: Alcohol intoxication facilitates creative problem solving. *Consciousness and Cognition*, *21*, 487–493.
- 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. *PLoS Biology*, *2*, 500–510.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.
- 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*, 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). New York, NY: Oxford University Press.
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47–70.
- 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*, 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*, 189–217.
- Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*, 348–354.

- Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of “executive attention.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 749–777.
- Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*, 374–419.
- Klinger, E. (1971). *Structure and functions of fantasy*. New York, NY: Wiley-Interscience.
- Klinger, E. (2013). Goal commitments and the content of thoughts and dreams: Basic principles. *Frontiers in Psychology*, *4*, 415. <http://journal.frontiersin.org/article/10.3389/fpsyg.2013.00415/abstract>
- König, C. J., Bühner, M., & Mürling, G. (2005). Working memory, fluid intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human Performance*, *18*, 243–266.
- Lee, C. S., & Theriault, 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*, 375–380.
- Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of Jazz improvisation. *PLoS One*, *3*, e1679. [10.1371/journal.pone.0001679](https://doi.org/10.1371/journal.pone.0001679)
- 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*, 85–96.
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, *52*, 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.
- McMillan, R. L., Kaufman, S. B., & Singer, J. L. (2013). Ode to positive-constructive daydreaming. *Frontiers in Psychology*, *4*, 626. <http://journal.frontiersin.org/article/10.3389/fpsyg.2013.00626/abstract>
- 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*, 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*, 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*, 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*, 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.
- Mednick, S. A., & Mednick, M. T. (1967). *Examiner’s manual: Remote Associates Test*. Boston, MA: Houghton Mifflin.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*, 806–834.
- Mullis, K. (1998). *Dancing naked in the mind field*. New York, NY: Pantheon Books.
- 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*, 36–45.
- Nusbaum, E. C., Silvia, P. J., & Beaty, R. E. (2014). Ready, set, create: What instructing people to “be creative” reveals about the meaning and mechanisms of divergent thinking. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 423–432. 10.1037/a0036549
- Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 61–65.
- Oberauer, K., Süß, H-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity—Facets of a cognitive ability construct. *Personality and Individual Differences*, *29*, 1017–1045.
- Oberauer, K., Süß, H-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167–193.
- Poole, B. J., & Kane, M. J. (2009). Working memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*, 1430–1454.
- Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin*, *140*, 1411–1431.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164–171.
- Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology, 20*, 713–721.
- Reverberi, C., Toraldo, A., D'Agostini, S., & Skrap, M. (2005). Better without (lateral) frontal cortex? Insight problems solved by frontal patients. *Brain: A Journal of Neurology, 128*, 2882–2890.
- Ricks, T. R., Turley-Ames, K. J., & Wiley, J. (2007). Effects of working memory capacity on mental set due to domain knowledge. *Memory & Cognition, 35*, 1456–1462.
- Sadowski, C. J., & Cogburn, H. E. (1997). Need for cognition in the big-five factor structure. *The Journal of Psychology: Interdisciplinary and Applied, 131*, 307–312.
- Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97–133). 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*, 319–326.
- Shute, V. J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research, 7*, 1–24.
- Silvia, P. J. (2008a). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences, 44*, 1012–1021.
- Silvia, P. J. (2008b). Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan (1965). *Creativity Research Journal, 20*, 34–39.
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., . . . 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*, 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.
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science, 26*, 559–569. [10.1177/0956797614567341](https://doi.org/10.1177/0956797614567341)
- Singer, J. L. (1966). *Daydreaming: An introduction to the experimental study of inner experience*. New York, NY: Random House.
- 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, NY: Academic Press, Inc.



- Singer, J. L., & Antrobus, J. S. (1963). A factor-analytic study of daydreaming and conceptually related cognitive and personality variables. *Perceptual and Motor Skills*, 17(Monograph Suppl. 3-V17), 187–209.
- Singer, J. L., & Schonbar, R. A. (1961). Correlates of daydreaming. *Journal of Consulting Psychology*, 25, 1–6.
- Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, 135, 94–120.
- Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process-occurrence framework for self-generated mental activity. *Psychological Bulletin*, 139, 519–535.
- 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, 527–533.
- Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, 36, 1144–1450.
- 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, 946–958.
- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66, 487–518.
- Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *The American Journal of Psychology*, 104, 61–87.
- Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic Bulletin & Review*, 14, 840–845.
- Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition*, 23, 551–559.
- Teasdale, J. D., Proctor, L., Lloyd, C. A., & Baddeley, A. D. (1993). Working memory and stimulus-independent thought: Effects of memory load and presentation rate. *European Journal of Cognitive Psychology*, 5, 417–433.
- Tierney, J. (2010, June 29). Discovering the virtues of a wandering mind. *The New York Times (Findings column)*. [http://www.nytimes.com/2010/06/29/science/29tier.html?ref=science&pagewanted=all&\\_r=0](http://www.nytimes.com/2010/06/29/science/29tier.html?ref=science&pagewanted=all&_r=0).

- Torrance, E. P. (2008). *Torrance Tests of Creative Thinking: Norms-technical manual, verbal forms A and B*. Bensenville, IL: Scholastic Testing Service.
- Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, *30*, 128–134.
- Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language*, *67*, 1–16.
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 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*, 832–842.
- Unsworth, N., McMillan, B. D., Brewer, G. A., & Spillers, G. J. (2012). Everyday attention failures: An individual differences investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1765–1772.
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent variable analysis of the relationship between processing and storage. *Memory*, *17*, 635–654.
- Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence*, *38*, 111–122.
- Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1302–1321.
- Unsworth, N., & Spillers, G. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*, 392–406.
- 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*, 289–335.
- Wiley, J., & Jarosz, A. F. (2012). How working memory capacity affects problem solving. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (pp. 185–228). New York, NY: Academic Press.

## APPENDICES

### APPENDIX A: WMC × Concentration Hierarchical Linear Models on Mind Wandering

<b>Experiment 1</b>					
<b>Fixed effect</b>	<b>Coefficient</b>	<b>SE</b>	<b>t ratio</b>	<b>df</b>	<b>p</b>
<i>Note.</i> WMC = working memory capacity; calculated by averaging across operation span and symmetry span z scores. Experiment 1 $N = 120$ , Experiment 2 $N = 115$ , Experiment 3 $N = 107$ . In these analyses, the main effects of WMC and concentration are represented by $\beta_{01}$ and $\beta_{10}$ , respectively, and the interaction between the two is represented by $\beta_{11}$ .					
Intercept, $\beta_{00}$	-.036	.122	-.292	118	.771
WMC, $\beta_{01}$	-.218	.135	-1.619	118	.108
Concentration, $\beta_{10}$	-1.044	.055	-18.908	13975	.000
WMC, $\beta_{11}$	.022	.067	.336	13975	.737
Random effect	$\chi^2$	<i>SD</i>	Variance component	<i>df</i>	<i>p</i>
Intercept, $r_0$	2599.634	1.325	1.755	118	.000
<b>Experiment 2</b>					
<b>Fixed effect</b>	<b>Coefficient</b>	<b>SE</b>	<b>t ratio</b>	<b>df</b>	<b>p</b>
Intercept, $\beta_{00}$	-.977	.172	-5.667	113	.000
WMC, $\beta_{01}$	-.050	.209	-.240	113	.811
Concentration, $\beta_{10}$	-1.100	.063	-17.549	17214	.000
WMC, $\beta_{11}$	-.109	.098	-1.115	17214	.265
Random effect	$\chi^2$	<i>SD</i>	Variance component	<i>df</i>	<i>p</i>
Intercept, $r_0$	4126.116	1.811	3.281	113	.000
<b>Experiment 3</b>					
<b>Fixed effect</b>	<b>Coefficient</b>	<b>SE</b>	<b>t ratio</b>	<b>df</b>	<b>p</b>
Intercept, $\beta_{00}$	.434	.179	2.427	105	.017
WMC, $\beta_{01}$	.063	.257	.246	105	.806
Concentration, $\beta_{10}$	-1.189	.106	-11.270	1180	<.001
WMC, $\beta_{11}$	-.017	.149	-.115	1180	.908
Random effect	$\chi^2$	<i>SD</i>	Variance component	<i>df</i>	<i>p</i>
Intercept, $r_0$	496.339	1.697	2.880	105	<.001

### APPENDIX B: *n*-back Stimuli for Experiment 1 Mind-Wandering Tasks

<b>Countries</b>	<b>Body parts</b>	<b>Instruments</b>	<b>Colors</b>
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 C: Experiment 1 Mean Accuracy and Reaction Times by Incubation Task

Task	Accuracy		RT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Note.</i> SART = sustained attention response task.				
SART target trials	.420	.222	369.469	77.352
SART nontarget trials	.933	.049	446.479	89.661
Coins <i>n</i> -back target trials	.772	.196	850.984	230.546
Coins <i>n</i> -back nontarget trials	.973	.046	1095.834	574.543
Coins <i>n</i> -back lure trials	.870	.108	981.211	298.798
Pigpen <i>n</i> -back target trials	.793	.186	843.432	229.570
Pigpen <i>n</i> -back nontarget trials	.977	.045	1010.872	464.521
Pigpen <i>n</i> -back lure trials	.877	.100	1058.591	321.313
Brick <i>n</i> -back target trials	.768	.215	840.427	229.886
Brick <i>n</i> -back nontarget trials	.970	.048	909.151	465.074
Brick <i>n</i> -back lure trials	.852	.112	1009.053	280.958
Knife <i>n</i> -back target trials	.765	.195	844.443	233.704
Knife <i>n</i> -back nontarget trials	.971	.045	984.665	519.659
Knife <i>n</i> -back lure trials	.876	.097	1053.134	367.940

### APPENDIX D: Experiment 2 Mean Accuracy and Reaction Times by Task

Task	Accuracy		RT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Note.</i> Reading comprehension questions were scored out of 5.				
Stand-alone reading task target trials	.440	.169	492.416	78.146
Stand-alone reading task nontarget trials	.981	.022	514.482	111.037
Coins reading task target trials	.521	.168	473.221	91.447
Coins reading task nontarget trials	.983	.024	487.750	103.516
Pigpen reading task target trials	.497	.179	480.617	99.992
Pigpen reading task nontarget trials	.983	.020	489.136	116.983
Brick reading task target trials	.484	.185	484.864	93.675
Brick reading task nontarget trials	.985	.014	463.303	111.890
Knife reading task target trials	.480	.172	484.397	97.585
Knife reading task nontarget 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	—	—

**APPENDIX E: Experiment 3 Mean Accuracy and Reaction Times for the Brick Incubation Task**

Task	Accuracy		RT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Note.</i> $N = 107$ for target trial accuracy, nontarget trial accuracy, and target trial reaction times; $N = 25$ for nontarget trial reaction times because of few responses during nontarget trials.				
Target trials	.878	.125	815.580	141.524
Nontarget trials	.997	.010	936.610	581.192