What are the benefits of mind wandering to creativity?

Samuel Murray[a]*, Nathan Liang[a], Nicholaus P. Brosowsky[b], and Paul Seli[a,b]
[a] Department of Psychology & Neuroscience, Duke University, Durham, North Carolina, USA
[b] Center for Cognitive Neuroscience, Duke University, Durham, North Carolina, USA

Author Note

Nicholaus P. https://orcid.org/0000-0003-4573-5468
Nathan Liang https://orcid.org/0000-0002-5308-6368
Samuel Murray https://orcid.org/0000-0002-4959-3252
Paul Seli https://orcid.org/0000-0002-5398-6999

Correspondence may be directed to Samuel Murray, Department of Psychology & Neuroscience, Duke University, 417 Chapel Drive, Durham, NC 27708-0086, USA.
Email: samuel.f.murray@duke.edu

Abstract

A primary aim of mind-wandering research has been to understand its influence on task performance. While this research has typically highlighted the *costs* of mind wandering, a handful of studies have suggested that mind wandering may be beneficial in certain situations. Perhaps the most-touted benefit is that mind wandering during a creative-incubation interval facilitates creative thinking. This finding has played a critical role in the development of accounts of the adaptive value of mind wandering and its functional role, as well as potential mechanisms of mind wandering. Thus, a demonstration of the replicability of this important finding is warranted. Here, we attempted to conceptually replicate results of a highly cited laboratory-based experiment supporting this finding. However, we found no evidence for the claim that mind wandering during a creative-incubation interval facilitates a form of creativity associated with divergent thinking. We suggest that our failed conceptual replication stems from an inadequate characterization of mind wandering (task-unrelated thought), and that there are good reasons to think that off-task thinking occurring in laboratory settings is unlikely to be causally related to creativity. Our results cast doubt on the claim that task-unrelated thought enhances creativity while also offering some prescriptions for how future research might further elucidate the cognitive benefits of mind wandering.

Keywords: Mind wandering; Task-Unrelated Thought; Incubation; Creativity; Divergent thinking

> "Don't think. Thinking is the enemy of creativity…You can't try to do things. You simply must do things."
> –*Ray Bradbury*

1. Introduction

What makes somebody or something creative? Typically, something—like an idea, product, or solution—is considered creative when it is new. The creative musician generates a unique combination of sounds, the creative chef discovers an original combination of flavors, and so on. But novelty isn't everything: according to many, the invention must exhibit some value to somebody (Dietrich, 2019; Hennessey & Amabile, 2009; Kenett et al., 2020). After all, any hack can throw together random ingredients from their kitchen. What sets the chef apart from the hack—true creativity—is the genesis of something that is both innovative and instrumental.

Different kinds of problem spaces require different kinds of creativity. Some problems have unique solutions, and in these spaces, the creative individual is one who can navigate the problem space to converge on this single solution. For example, the mathematician building a proof is searching for a single conclusion. This is a paradigmatic example of *convergent* creativity. Other problems require generating numerous potential solutions, and the creative individual is judged to be the one who can generate many unique solutions. For example, the screenwriter searching for new story ideas benefits from being able to flexibly generate a sizeable list of narratives. This is an exemplar of *divergent* creativity.

Creativity is crucial for many of the unique and valuable endeavors we undertake. It's important and often necessary for facilitating scientific, artistic, and political achievements that underscore the heights of human progress. But creativity also plays a role in mundane interactions, from improvising a recipe when lacking ingredients to maintaining engaging conversations over dinner. Creative people report experiencing better mood (Nadler et al., 2010), higher self-esteem (Barbot, 2018), and score higher on several dimensions of well-being (Conner et al., 2016). However, despite its demonstrated importance and benefits, insight into methods of facilitating or enhancing creativity remains sorely lacking.

Anecdotal evidence about the mannerisms and discoveries of highly creative individuals provide some clues about the underlying processes supporting creative cognition. Many creative insights seem to occur in moments wherein the person is not actively focusing on generating a creative insight. Kekulé, for example, claims to have discovered the ring structure of the benzene molecule in a dream. The Indian mathematician Ramanujan was struck with numerous thoughts, the source of which he could not explain (so much so that he attributed his insights to a local Hindu goddess of good fortune; Cheng, 2017). Poincaré, Einstein, and Edison, among many others, all made similar claims. The collection of such anecdotal examples suggests an interesting potential phenomenon: setting a problem aside and not consciously thinking about it can lead to creative breakthroughs. This temporary disengagement is known as *incubation*, and there is a wealth of evidence to suggest that it is beneficial to creative thinking. For example, Gable et al. (2019) recently found that, when asked to report the creative quality and context of their most salient daily "aha" moments, physicists and writers exhibited a pattern of more frequently generating enlightening solutions during mind wandering relative to focused thinking.

The idea that incubation is readily conducive to creativity invites the question of what cognitive modes might promote incubation in the first place. One relevant mode seems to be mind wandering. Consider that incubation requires setting a problem aside and thinking about something else. Mind wandering—as it's most frequently characterized—reflects disengaging with a focal task to process internally generated (task-irrelevant) contents (Smallwood & Schooler, 2015). Thus, when someone is mind wandering, they are thinking about something unrelated to their focal task. Mind wandering, then, constitutes a form of distraction which might invite incubation. Altogether, this implies that mind wandering should enhance creativity by promoting incubation.

One influential study found evidence supporting this conjecture: Baird et al. (2012) examined performance on the Alternate Uses Task (AUT; a widely used measure of divergent creativity; Guilford, 1967) under several conditions imposed following participants' initial efforts to think of novel and creative uses for a common object (e.g., brick). These conditions included: (a) performing a demanding (2-back) task, (b) performing a less-demanding (0-back) task, (c) having a period of rest, or (d) performing an immediate repetition of the AUT, with no intermediate activity. The authors hypothesized that the less-demanding task would produce more mind wandering than the demanding task, and that this would in turn lead to improved performance on the AUT (interpreted as greater creativity). Consistent with their hypotheses, the less-demanding task was found to be associated with more mind wandering than the demanding task and, critically, the less-demanding task was also associated with greater subsequent production of additional creative 'uses' responses relative to the demanding task. Thus, the authors concluded that mind wandering led to the discovery of new uses beyond those imagined in the first exposure.

This finding represented a major advance in the study of both creativity and mind wandering (as evidenced by the 854 citations Baird et al.'s paper has garnered as of September 30, 2020). On the creativity side, this presented a promising intervention to facilitate creative-idea generation. Additionally, since some mind wandering is intentional, if mind wandering enhances creativity, people would be able to deliberately cultivate mindsets that boost their creativity. In the mind-wandering camp, this finding provided evidence for speculative accounts of the adaptive value and functional role of mind wandering, and it contributed to the development of novel frameworks of mind wandering.

Smeekens and Kane (2016) ran a series of conceptual replications of Baird et al. (2012). Across several multi-session studies measuring how working-memory capacity and other individual-differences variables affected mind wandering and creativity, Smeekens and Kane failed to find the expected significant relationships between mind wandering and AUT performance, despite utilizing various scoring procedures. In their Experiment 3 (the one modelled most closely on Baird et al.), they had participants perform the AUT, followed by an undemanding incubation-interval task, followed by an additional repetition of the AUT. Even here, Smeekens and Kane (2016) failed to find any relation between mind wandering during the incubation-interval and AUT scores.[1]

---

[1] Notably, Leszczynski et al. (2017) found that mind wandering positively correlates with creative incubation. Participants completed the SART between iterations of the Compound Remote Associates Test (Lee, Huggins, & Therriault, 2014). In the latter, participants are given three words and asked to find the shared associate (e.g., "cottage—Swiss—cake" would have "cheese" as its solution). Participants who reported higher proportions of off-task thought during the SART also produced more solutions in the second round of the task relative to those who reported lower proportions of off-task thought. However, unlike Baird et al. (2012), Leszczynski et al. used a test of *convergent* creativity rather than *divergent* creativity. Hence, these results cannot be considered a replication of Baird et al. (2012).

We present here the results of a conceptual replication of Baird et al. (2012) with some important differences between our experiment and those of Smeekens and Kane (2016). We adopted three design features from Smeekens and Kane. First, we explicitly prompted participants to generate creative and useful responses to the AUT. Second, Baird et al. assessed AUT performance using uniqueness scores, where unique responses (relative to the entire set) are assigned a score of 1 (non-unique responses are assigned a score of 0). However, given criticisms of uniqueness scoring for measuring creativity (Silvia et al., 2009; Smeekens and Kane, 2016), we used subjective ratings of creativity generated by three independent raters. Finally, we measured mind wandering using thought probes rather than a retrospective report. Baird et al. measured mind wandering with a single retrospective report collected at the end of the incubation-interval. However, this method is likely to be less reliable than *in situ* probes because accurately responding to a single retrospective measure requires significantly more working memory and recollection than accurately responding to a probe about where attention was directed over the previous few seconds.

Our study more closely conceptually replicates Baird et al. (2012) relative to Smeekens and Kane (2016) for several reasons. First, Smeekens and Kane did not manipulate proportions of mind wandering during the incubation interval and measure for effects of this manipulation; instead, they used a single-condition study looking at predictive relations between proportion of mind wandering during an undemanding incubation-interval task and AUT responses. In this experiment, we use a task-difficulty manipulation from Baird et al. to manipulate proportion of mind wandering across conditions. Second, Smeekens and Kane did not measure the content of off-task thoughts during the incubation-interval task. This is important, as thoughts about the AUT during the incubation interval would count as mind wandering according to the probes used by Smeekens and Kane because such thoughts are unrelated to the incubation interval task. However, such thoughts would *not* constitute incubation, as incubation about some non-focal task or problem precludes occurrently thinking about the task or problem. Neither of these are meant to indicate problems with Smeekens and Kane, as their studies examined more general relationships between executive control, working memory, and different forms of creativity. Hence, they did not set out to replicate the results from Baird et al. (2012). Additionally, unlike either Baird et al. or Smeekens and Kane, we (a) collected data from a larger sample, (b) explicitly informed participants about the post-incubation AUT, which should encourage incubation effects, and; (c) used Bayesian analyses to supplement significance testing.

2. Method

Pre-registration of sample size, primary outcome measures, exclusion criteria, experimental materials (including experiment programs and stimuli), raw data, and analysis scripts can be found at https://osf.io/dwec2/. In accordance with the recommendations of Simmons et al. (2012), we report how we determined our sample size, all data exclusions, and all measures in our study. All procedures were approved by the Duke University Internal Review Board.

2.1 Participants

We recruited 200 participants (M_{age} = 40.24, SD_{age} = 11.22, female = 104) through Amazon's Mechanical Turk. We recruited only participants located in the United States with a HIT approval rate greater than or equal to 98% and at least 5000 previously approved HITs. We determined our target sample size based on the results of a priori power analyses conducted in the G*Power 3.1.9.7 software (Faul et al., 2007) for independent-samples *t*-tests with a medium effect size (*d* = 0.40), standard two-tailed alpha value (*p* < .05), and an acceptable power threshold (1- β err prob = 0.80). Per the criteria described in the pre-registration, we excluded data from six participants based on their self-reported use of online resources while completing the AUT (final *N* = 194).

2.2 Materials

2.2.1 N-back Task

We induced differential proportions of mind wandering by implementing a between-groups n-back manipulation (Smallwood et al., 2011; Baird et al., 2012; Konishi et al., 2015; Smeekens & Kane, 2016; Seli, Konishi, Risko, & Smilek, 2018; Brosowsky et al., under review). In this paradigm, participants are asked to respond to target stimuli and withhold responses to non-target stimuli. The stimuli consisted of eight digits (1-8) presented serially on-screen, with target stimuli displayed in red and non-target stimuli displayed in black. In the *0-back condition*, participants were asked to indicate whether the red digit was even or odd. In the *2-back condition*, participants were asked to indicate whether the digit presented two trials prior to the red digit was even or odd.

At the start of each trial, a fixation cross was presented in the center of the screen for 1500ms followed by a blank screen presented for 500ms. The target stimulus was displayed on screen for 1500ms followed by a blank screen for 500ms. If a response was not registered within the 1500ms window, the trial was counted as a miss. Participants completed one practice block containing 24 trials with 4 target stimuli. If participants responded to a non-target (false alarm), they were instructed to withhold responses to non-targets. If participants did not respond or responded incorrectly to a target (miss), they were instructed to respond according to the instructions. Participants had to respond correctly to 3 out of 4 targets to move on to the experimental trials. Participants were not given feedback on performance in the experimental trials.

2.2.2 Thought Probes

At semirandom intervals, participants were presented with thought probes to assess the content of thought (Smallwood & Schooler, 2015; Weinstein, 2018). Probes were never presented immediately after a target stimulus. A thought-probe trial was presented once in every 16 trials (15 n-back trials/1 thought-probe trial), randomly presented between trials 7 to 11. The minimum time between probes was 48 seconds and the maximum time between probes was 80 seconds (64-second intervals, on average). These probes inquired about what the participant was thinking about just prior to seeing the probe with three options: (1) thinking about the even/odd task; (2) thinking about the upcoming creativity task, and; (3) thinking about something unrelated to the experiment. The second option was included because participants were explicitly instructed about the AUT to be completed. Reaction times were recorded for each thought probe response.

In keeping with the proposal from Seli et al. (2018), and the experimental procedures from Baird et al. (2012), we operationalized mind wandering in terms of task-unrelated thought. Thus, when participants reported either thinking about the upcoming creativity task or thinking about something unrelated to the experiment, these were recorded as mind wandering. This matches the procedures used in Baird et al. (2012) and Smeekens & Kane (2016).

2.2.3 Alternate Uses Task

The Alternate Uses Task (AUT; Guilford, 1967) is a prominent measure of divergent creativity (Plucker & Makel, 2010) that assesses individual ability to access semantically distant concepts relative to a mundane cue. Participants were provided the name of a single everyday object ("brick") and asked to generate creative and unusual uses for the object. Three sample responses were provided. Additionally, participants were told to generate responses that are creative, useful, and specific to the object.

2.3 Procedure

After reading instructions on the n-back task, AUT, and thought probes, participants practiced the n-back task and could not advance without responding correctly to three targets. Before beginning the experimental trials, participants were again told about completing the AUT immediately after the n-back. Participants completed 15 minutes of n-back trials before doing the AUT for 2.5 minutes. A timer was visible for the AUT, indicating how much time was remaining (no timer was visible for the n-back task).

2.4 Creativity Scoring

Three raters (undergraduate research assistants trained in AUT scoring but blind to the hypotheses of the study and conditions) independently scored each participant's individual responses on a scale of 1-5. To assign scores, raters were told to assess responses on their novelty and creativity and to exclude nonsense answers with the following scoring system (full instructions are available at https://osf.io/dwec2/): 1 = very obvious/ordinary use; 2 = somewhat obvious use; 3 = non-obvious use; 4 = somewhat imaginative use; 5 = very imaginative/re-contextualized use; 99 = invalid (all invalid responses were excluded from final analyses). Raters exhibited strong reliability ($\alpha = 0.906$). Creativity scores were calculated by taking the arithmetic mean of all three ratings.

3. Results

We supplemented null hypothesis significance tests with Bayes Factor analyses using the BayesFactor package in R with default settings (Morey & Rouder, 2018). To simplify interpretation, we report the Bayes Factor in the direction the data supports (BF_{01} when there is more evidence in favor of the null over alternative hypothesis and BF_{10} when there is more evidence in favor of the alternative over null hypothesis).

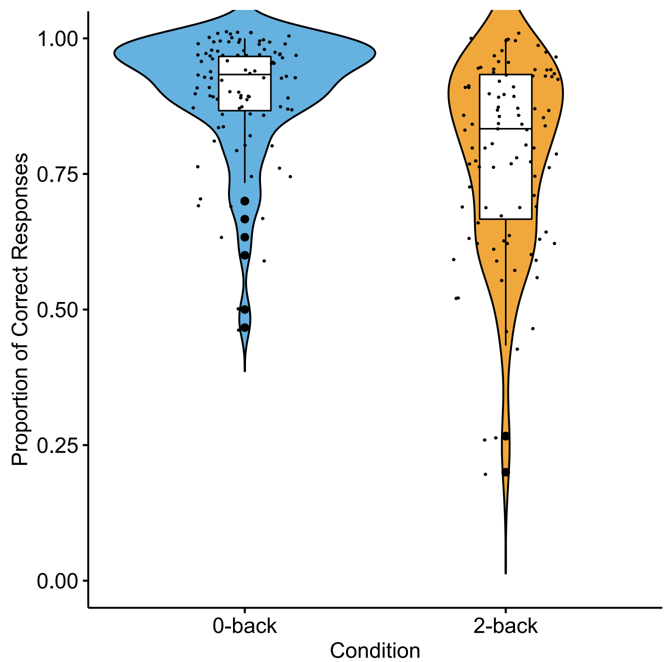

Figure 1. Proportion of correct responses between n-back conditions (0-back or 2-back)

3.1 Task difficulty manipulation

To assess the effectiveness of the task difficulty manipulation, we compared performance and proportion of mind wandering across conditions. Based on previous studies, we hypothesized that participants in the 0-back condition would perform better and mind wander more than participants in the 2-back condition. Performance was calculated as the proportion of correct responses to total responses. Participants in the 0-back condition performed significantly better (M = .902, SD = 0.110) than participants in the 2-back (M = .782, SD = 0.173) condition (Welch's $t(162) = 5.77$, $p < .001$), with overwhelming evidence in favor of the alternative hypothesis over the null, $BF_{10} = 3.52 \text{x} 10^5$ (Figure 1).

As expected, participants reported significantly more mind wandering in the 0-back (37.1% of thought probes) relative to the 2-back (17.1% of thought probes) condition (Welch's $t(161) = 5.78$, $p < .001$), with overwhelming evidence in favor of the alternative hypothesis $BF_{10} = 3.73 \text{ x } 10^5$ (Figure 2). Both performance and wind wandering measures resemble results found in other studies utilizing task difficulty manipulations to influence rates of mind wandering (Baird et al., 2011; Brosowsky et al., under review; Smeekens & Kane, 2016).

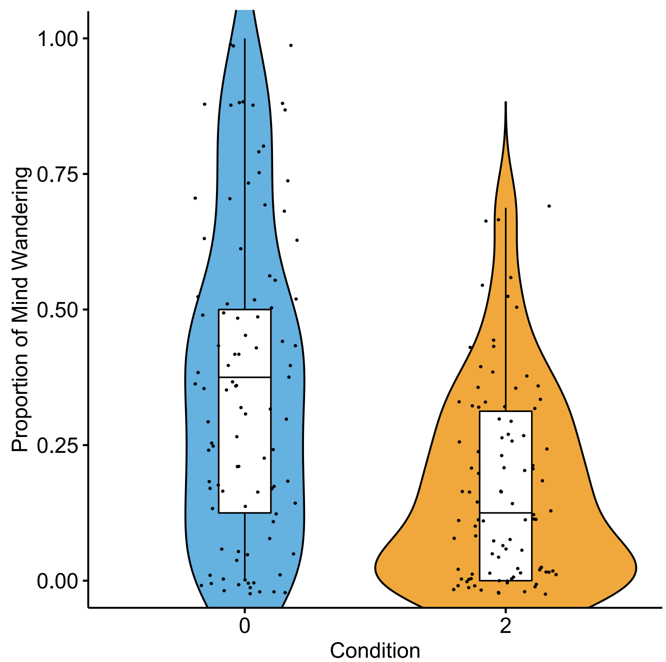

Figure 2. Proportion of mind wandering between n-back conditions (0-back or 2-back)

3.2 The effect of mind wandering on creativity

To assess whether mind wandering is associated with AUT scores, we compared AUT performance across conditions. An independent samples *t*-test did not yield significant differences in AUT scores across conditions (Student's $t(191) = 1.54$, $p = .125$; cf. Figure 3) and Bayesian analyses yielded modest evidence in favor of the null hypothesis $BF_{01} = 2.11$ and anecdotal evidence in favor of the alternative hypothesis, $BF_{01} = 0.473$.

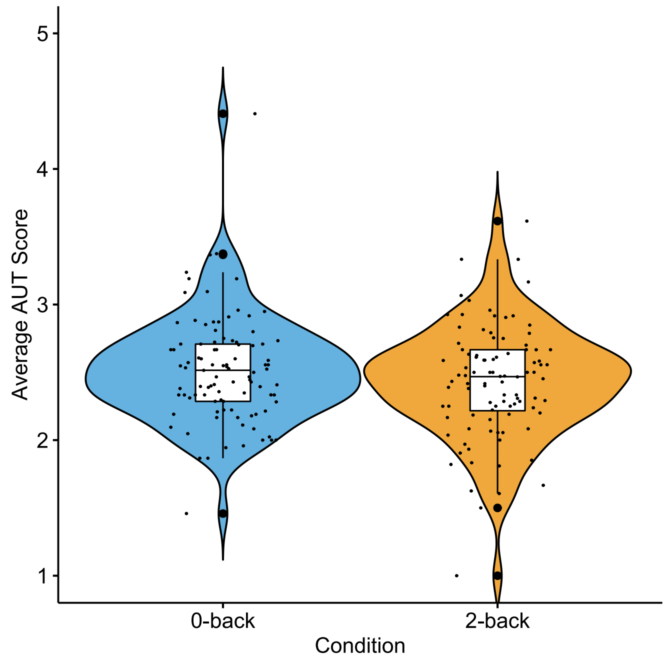

Figure 3. Average AUT scores (1-5) between n-back conditions (0-back and 2-back).

Moreover, collapsing across conditions, we failed to find a significant correlation between individual proportion of mind wandering and average AUT score (r = -0.038, p = 0.595), with strong evidence in favor of the null hypothesis, BF_{01} = 9.65 (cf. Figure 4).

As in Baird et al. (2012), we split participants' data according to whether they registered at least one AUT-related thought during the n-back task or not. In the 0-back condition, 52 participants registered at least one AUT-related thought (45 participants registered none), whereas in the 2-back condition, 40 participants registered at least one AUT-related thought (56 participants registered none). AUT scores were not significantly different between those who had at least one AUT-related thought (M = 2.47, SD = 0.41) and those with none (M = 2.48, SD = 0.41), $t(191)$ = 0.148, p = 0.883 with anecdotal evidence in favor of the hypothesis that AUT scores differed as a function of AUT-related thought, BF_{10} = 0.775. We also submitted AUT scores to a between-subjects ANOVA with n-back group and AUT-related thought as factors, which revealed no significant effect of condition ($F(1, 189)$ = .01, MSE = .01, p = 0.748) with moderate evidence for the null hypothesis over the hypothesis that there is an effect of AUT-related thoughts on AUT score, BF_{01} = 6.32.

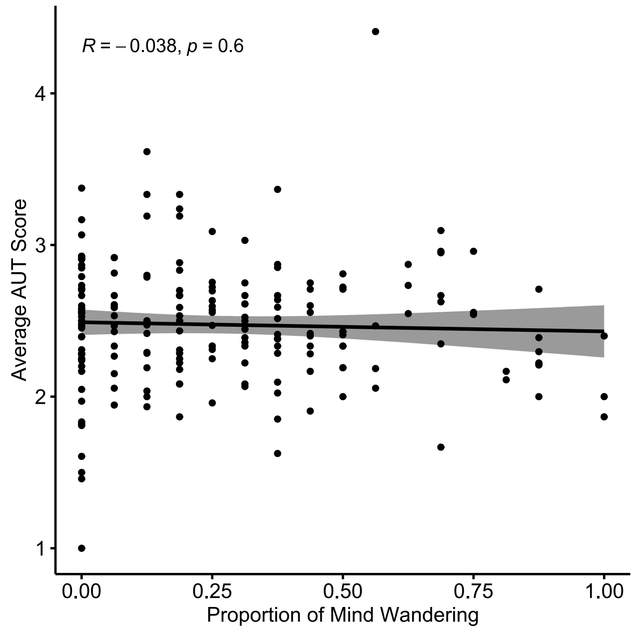

Figure 4. Correlation between average AUT score and individual proportion of mind wandering.

3.3 Exploratory Analyses

Finally, for exploratory purposes, we used a recently developed automated AUT scoring program (SemDis; Beaty & Johnson, 2020) to assess whether the predicted effect of mind wandering on creativity obtains when using automated creativity scores. SemDis computes the semantic distance between the cue and the participant response relative to a semantic space. The program enables computing semantic distance within five different semantic spaces (cbowukwacsubtitle_nf_m, cbowsubtitle_nf_m, cbowBNCwikiukwac_nf_m, TASA_nf_m, and glove_nf_m). We chose to analyze responses using automated ratings indexed to each of the five semantic spaces, as well as a composite score based on all five sets of ratings (SemDis_MEAN). We first wanted to assess correlations between subjective ratings and automated ratings. Because subjective and automated ratings are generated on different scales, we first normalized both sets of ratings using Min-Max scaling with the scikit-learn package in Python. This maps all values onto a common [0, 1] range. After normalization, a Pearson product-moment correlation revealed a moderate correlation between subjective scoring and automated scoring (r = .350, p < .001). Independent-samples t-tests revealed no significant differences and strong evidence in favor of the null hypothesis in creativity scores across conditions for all five semantic spaces, as well as composite mean and factor scores (see Table 2).

Table 2. Bayesian analyses of automatically scored AUT responses across n-back conditions using different semantic spaces.

	BF_{0+}	error %
SemDis_cbowukwacsubtitle_nf_m	63.7	0.00435
SemDis_cbowsubtitle_nf_m	25.3	0.18666
SemDis_cbowBNCwikiukwac_nf_m	48.2	0.11387
SemDis_TASA_nf_m	12.3	1.72e-6
SemDis_glove_nf_m	40.0	0.02815
SemDis_MEAN	47.3	0.02769
SemDis_factor	51.4	0.01042

Note. For all tests, the alternative hypothesis specifies that group *0-back* is greater than group *2-back*.

4. Discussion

Here, we sought to replicate results reported in Baird et al. (2012) that an increased proportion of mind wandering during an incubation interval correlates with better performance on a creativity task that recruits divergent thinking. We replicated Baird et al.'s finding that task-difficulty manipulations modulate rates of mind wandering: harder tasks tend to elicit less overall mind wandering than easier tasks. However, unlike Baird et al., we did not find evidence for an effect of condition on creativity. Moreover, we did not find that individuals' proportions of mind wandering predicted their creativity scores.

These results align with the failed conceptual replication reported in Smeekens and Kane (2016). However, the latter did not measure the content of mind wandering during the incubation interval. Hence, Smeekens and Kane could not confirm that the n-back task functioned as an incubation period for the creativity task because they could not rule out participants explicitly thinking about the creativity task during the n-back. In this study, we split participants' responses according to whether they thought of the subsequent creativity task at all during n-back performance or not. This allowed us to remove people who are mind wandering in ways that fail to promote incubation.

Even after splitting the data, we did not find any significant relationships between mind wandering and AUT performance. This might seem surprising in one of two directions, so it's important to understand what this result indicates. On the one hand, isolating the people who did not think about the AUT at all during the n-back picks out people for whom the n-back might function as an incubation interval. Thus, we should expect that isolating the potential incubators in the undemanding condition would generate the predicted effect of mind wandering on creativity. However, even after making these post-hoc distinctions, we did not find that individual proportion of mind wandering predicts subsequent AUT performance. On the other hand, it might seem surprising that people who report thinking about the AUT during the n-back do not perform better on it relative to those who don't. While this result appears surprising, it should be interpreted cautiously. We did not account how much time one spent thinking about the AUT, so measure of AUT-related thought might not be fine-grained enough to reveal an effect of thinking on creativity performance (however, because the number of AUT thoughts was so low to begin with, it seems unlikely that a more sophisticated tool would yield different results). This result might also not be as surprising as it seems at first blush. Perseverative thinking might strengthen a narrow range of associations, resulting in a person producing a constrained range of AUT responses. Mind wandering is known to be associated with hippocampally-mediated episodic memory reactivation (Ellamil et al., 2016; McCormick et al., 2018), which can lead to distinct contents being reactivated jointly in a novel context. This can facilitate forming novel associations between contents that are less likely to form during periods of goal-directed thinking (Mills et al., 2018). Thus, there is some evidence that more thinking about the AUT would not necessarily result in better performance on it.

The effects of incubation on creativity are well-known (cf. Gilhooly, 2016). Incubation requires that one refrain from consciously attending to a problem, and during such periods, bouts of unconscious processing serve to facilitate insights and breakthroughs with respect to some problem. Given that mind wandering consists, partly, in off-task thinking, why do we fail to find any effects of mind wandering on creativity via incubation?

One issue concerns the theoretical operationalization of mind wandering as task-unrelated thought. Baird et al. (2012), following standard procedures in research on mind wandering, assessed mind wandering in terms of thoughts unrelated to a focal task. However, this category is too heterogeneous to function as a useful tool for measuring mind wandering. Each term in the operationalization is ambiguous. It's unclear what makes something a task and at what level of behavior we describe tasks. For example, is earning a degree a task? Or is earning a degree a goal, with writing papers as a task? Or should we count pen strokes and key strikes as tasks that accumulate into paper writing? If these are all different kinds of tasks, then what kind of task is relevant for assessing mind wandering? Moreover, what makes a thought *related* to a task? Suppose you are performing a SART and asked to respond only to red targets. The first red target you see reminds you of strawberries you ate for breakfast, and you begin thinking about when you might be able to buy more. This thought is related to the SART in some ways but not others. What is the kind of relation relevant to measuring mind wandering? The standard operationalization settles none of these issues, as well as a host of additional ambiguities (see (Murray et al., 2020 for further discussion). Thus, too many kinds of thoughts count as task-unrelated in certain contexts. And, there's no reason to expect that all thoughts falling under this wide umbrella would uniformly facilitate creativity.

This heterogeneity problem arises in connection with the current replication and the original studies. During the n-back task, some participants might think about the AUT. According to standard methods, these thoughts count as mind wandering, but they don't count as incubation. In fact, task-unrelated thoughts do not indicate anything about incubation unless one first has reason to believe that most task-unrelated thoughts do not gravitate toward thinking about planned behaviors. In that case, one could say that when we observe TUTs these are most likely (though not definitely) constitutive of incubation. But a wealth of evidence indicates precisely the opposite: mind wandering has a general prospective orientation and often takes planned behaviors as part of its content (Baird et al., 2011). Thus, task-unrelated thought is too variable and wide-ranging to reliably constitute thinking that facilitates creativity.

Incubation requires unconscious processing of goal-relevant information, which enables one to make progress on a stalled goal by exploring novel associations among mnemonic contents that are unlikely to obtain when consciously focusing on one's representation of the problem. Task-unrelated thought, however, must be either conscious or hopelessly heterogeneous (as argued above). Consider, first, that measurements of task-unrelated thinking typically require participants to report whether they are thinking about something other than some researcher-assigned task. This reporting uses some process of retrospective evaluation of previous thought content (typically over a short period of time) prompted by a thought probe. The assumption is that task-unrelated thinking—or the contents that warrant an inference that one was mind wandering—are conscious in the sense of being available for verbal report. The accessibility of this content implies that task-unrelated thinking is conscious in a way that precludes incubation.

Sometimes, researchers rely on overt behavioral cues or physiological signals to infer mind wandering to avoid using self-report to measure mind wandering (cf. Murray et al., Forthcoming), including reaction time variability (Bastian & Sackur, 2013; Seli et al., 2013), gaze features (Faber et al., 2020; Feng et al., 2013; Forrin et al., 2019; Uzzaman & Joordens, 2011), performance (McVay & Kane, 2009; Schooler et al., 2004), and even fidgeting (Seli et al., 2014). One issue with these purportedly objective measures of mind wandering is that they are validated against self-report measures of mind wandering. Measures of mind wandering typically characterize it in terms of task-unrelated thought, so these objective measures end up being little more than automated means of indexing task-unrelated thought. So, the problems mentioned previously about the compatibility of task-unrelated thought and incubation remain. Additionally, a crucial assumption underlying the inference from performance deficits to mind wandering is that task-irrelevant thoughts are detrimental to performance. Part of what makes a thought task-*unrelated* is this negative relationship to performance. While this kicks out the ladder of self-report, it leads directly back to the problem of heterogeneity: any number of thoughts can result in performance deficits on a focal task, and some mind wandering might not hinder performance of certain tasks (Brosowsky et al., revision submitted; Seli et al., 2018; Thomson et al., 2014). Thus, either task-unrelated thought is conscious, or it is inferred from assumptions that make task-unrelated thought empirically irrelevant to incubation.

As noted in the Introduction, the connection between mind wandering and creativity plays an important role in views about the functional role and adaptive value of mind wandering. Because the study from Baird et al. (2012) was the first to experimentally manipulate mind wandering to modulate creativity, it became a cornerstone of such theoretical accounts. Sripada (2018, p. 25), for instance, cites only Baird et al. (2012) to support his claim that mind wandering is an exploratory mode of cognition. Because of this failed replication (coupled with Smeekens and Kane, 2016), must we cede that these theoretical accounts are misguided?

We think that conclusion would be too rash. However, we do think this calls for a more nuanced methodological approach to studying the effects of mind wandering on creativity. One important methodological change is to reconsider the kind of creative benefits mind wandering is likely to generate. Potential effects of mind wandering on creativity are likely to appear for problems with some personal relevance to the individual (cf. Klinger, 2013). Hence, studies of mind wandering and creativity might consider using the Personal Concerns Inventory (Cox & Klinger, 1988) to assess areas of concern where people might experience creative breakthroughs. Second, researchers should reconsider the temporal scale along which effects of mind wandering on creativity are likely to occur. For Baird et al. (2012), mind wandering was supposed to influence creativity on an impersonal word association task within 15 minutes. However, breakthroughs on personally relevant problems or issues might require hours, days, or even weeks to occur. Hence, rather than using impersonal tests of creativity over short time scales, we suggest understanding creativity applied to the personal concerns of individuals over longer timescales. For this reason, we find the methodology of Gable et al. (2019) exemplary. They used experience-sampling techniques to assess the context, phenomenology, and content of thought during creative breakthroughs and how these moments of creativity related to the overall project. While this study focused on the creative insights of professional scientists and writers, it is also possible to use experience-sampling methods to assess the relationship between creativity, mind wandering, and everyday concerns. We see no reason to think that mind wandering plays a fundamentally different role in the psychic economy of academics than non-academics.

5. Conclusions

The results presented here indicate that mind wandering does not facilitate creative thinking by promoting incubation. Ultimately, the failure to replicate Baird et al. (2012), coupled with the failed replications reported in Smeekens and Kane (2016), indicate that the results should no longer be used to support theories of mind wandering or hypotheses about mind wandering. As indicated in the Discussion, this makes good sense. There are good reasons to think that the standard characterization of mind wandering (task-unrelated thought) used by Baird et al. is unlikely to bear interesting relationships to creativity and incubation. In this way, the failure to replicate is unsurprising.

While alternative methods, outlined in the Discussion, might reveal interesting relationships between creativity and mind wandering, that's a matter of what evidence there might be, not what evidence there is. This point bears repeating. Since 2017, the failed replications in Smeekens and Kane (2016) have been cited 68 times. Meanwhile, Baird et al. have steadily *increased* their citation rate, garnering over half of their citations (478) since 2017. We hope that our efforts induce some restraint among researchers looking to use the results from Baird et al.

6. Acknowledgements

This research was supported by a Charles Lafitte Foundation Undergraduate Research Grant to Nathan Liang through the Duke University Department of Psychology & Neuroscience.

References

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*. https://doi.org/10.1177/0956797612446024

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition*, *20*(4), 1604–1611. https://doi.org/10.1016/j.concog.2011.08.007

Barbot, B. (2018). The Dynamics of Creative Ideation: Introducing a New Assessment Paradigm. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02529

Bastian, M., & Sackur, J. (2013). Mind wandering at the fingertips: automatic parsing of subjective states based on response time variability. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00573

Beaty, R., & Johnson, D. R. (2020). *Automating Creativity Assessment with SemDis: An Open Platform for Computing Semantic Distance*. https://doi.org/10.31234/osf.io/nwvps

Brosowsky, N., Murray, S., Schooler, J. W., & Seli, P. (n.d.-a). *Attention need not always apply: Mind wandering impedes explicit but not implicit sequence learning*. 28.

Brosowsky, N., Murray, S., Schooler, J. W., & Seli, P. (n.d.-b). *Thought dynamics under task demands: Evaluating the influence of task difficulty on unconstrained thought*.

Cheng, E. (2017). *Beyond Infinity: An Expedition to the Outer Limits of Mathematics* (1st Edition). Basic Books.

Conner, T. S., DeYoung, C. G., & Silvia, P. J. (2016). Everyday creative activity as a path to flourishing. *The Journal of Positive Psychology*, *13*(2), 181–189. https://doi.org/10.1080/17439760.2016.1257049

Cox, W. M., & Klinger, E. (1988). A motivational model of alcohol use. *Journal of Abnormal Psychology*, *97*(2), 168–180. https://doi.org/10.1037/0021-843X.97.2.168

Dietrich, A. (2019). Types of creativity. *Psychonomic Bulletin & Review*, *26*(1), 1–12. https://doi.org/10.3758/s13423-018-1517-7

Ellamil, M., Fox, K. C. R., Dixon, M. L., Pritchard, S., Todd, R. M., Thompson, E., & Christoff, K. (2016). Dynamics of neural recruitment surrounding the spontaneous arising of thoughts in experienced mindfulness practitioners. *NeuroImage*, *136*, 186–196. https://doi.org/10.1016/j.neuroimage.2016.04.034

Faber, M., Krasich, K., Bixler, R. E., Brockmole, J. R., & D'Mello, S. K. (2020). The eye–mind wandering link: Identifying gaze indices of mind wandering across tasks. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000743

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review*, *20*(3), 586–592. https://doi.org/10.3758/s13423-012-0367-y

Forrin, N. D., Risko, E. F., & Smilek, D. (2019). On the relation between reading difficulty and mind-wandering: a section-length account. *Psychological Research*, *83*(3), 485–497. https://doi.org/10.1007/s00426-017-0936-9

Gable, S. L., Hopper, E. A., & Schooler, J. W. (2019). When the Muses Strike: Creative Ideas of Physicists and Writers Routinely Occur During Mind Wandering. *Psychological Science*, *30*(3), 396–404. https://doi.org/10.1177/0956797618820626

Gilhooly, K. J. (2016). Incubation and Intuition in Creative Problem Solving. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01076

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

Hennessey, B. A., & Amabile, T. M. (2009). Creativity. *Annual Review of Psychology*, *61*(1), 569–598. https://doi.org/10.1146/annurev.psych.093008.100416

Kenett, Y. N., Kraemer, D. J. M., Alfred, K. L., Colaizzi, G. A., Cortes, R. A., & Green, A. E. (2020). Developing a neurally informed ontology of creativity measurement. *NeuroImage*, *221*, 117166. https://doi.org/10.1016/j.neuroimage.2020.117166

Klinger, E. (2013). Goal Commitments and the content of thoughts and dreams: basic principles. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00415

McCormick, C., Rosenthal, C. R., Miller, T. D., & Maguire, E. A. (2018). Mind-Wandering in People with Hippocampal Damage. *Journal of Neuroscience*, *38*(11), 2745–2754. https://doi.org/10.1523/JNEUROSCI.1812-17.2018

McVay, J. C., & Kane, M. J. (2009). Conducting the Train of Thought: Working Memory Capacity, Goal Neglect, and Mind Wandering in an Executive-Control Task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(1), 196–204. https://doi.org/10.1037/a0014104

Mills, C., Herrera-Bennett, A., Faber, M., & Christoff, K. (2018, May 29). *Why the Mind Wanders*. The Oxford Handbook of Spontaneous Thought. https://doi.org/10.1093/oxfordhb/9780190464745.013.42

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for Common Designs* (0.9.12-4.2) [Computer software]. https://CRAN.R-project.org/package=BayesFactor

Murray, S., Irving, Z., & Krasich, K. (2020). *The Scientific Study of Passive Thinking: The Methodology of Mind Wandering Research* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/aqvrc

Nadler, R. T., Rabi, R., & Minda, J. P. (2010). Better Mood and Better Performance: Learning Rule-Described Categories Is Enhanced by Positive Mood. *Psychological Science*. https://doi.org/10.1177/0956797610387441

Plucker, J. A., & Makel, M. C. (2010). Assessment of creativity. In *The Cambridge handbook of creativity* (pp. 48–73). Cambridge University Press. https://doi.org/10.1017/CBO9780511763205.005

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning Out while Reading: Evidence for Dissociations between Experience and Metaconsciousness. In *Thinking and seeing: Visual metacognition in adults and children* (pp. 203–226). MIT Press.

Seli, P., Beaty, R. E., Cheyne, J. A., Smilek, D., Oakman, J., & Schacter, D. L. (2018). How pervasive is mind wandering, really? *Consciousness and Cognition*, *66*, 74–78. https://doi.org/10.1016/j.concog.2018.10.002

Seli, P., Carriere, J. S. A., Thomson, D. R., Cheyne, J. A., Martens, K. A. E., & Smilek, D. (2014). Restless mind, restless body. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(3), 660–668. https://doi.org/10.1037/a0035260

Seli, P., Cheyne, J. A., & Smilek, D. (2013). Wandering minds and wavering rhythms: linking mind wandering and behavioral variability. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(1), 1–5. https://doi.org/10.1037/a0030954

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-Wandering With and Without Intention. *Trends in Cognitive Sciences*, *20*(8), 605–617. https://doi.org/10.1016/j.tics.2016.05.010

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 Word Solution* (SSRN Scholarly Paper ID 2160588). Social Science Research Network. https://doi.org/10.2139/ssrn.2160588

Smallwood, J., & Schooler, J. W. (2015). The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annual Review of Psychology*, *66*(1), 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

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. https://doi.org/10.1037/aca0000046

Sripada, C. S. (2018, May 29). *An Exploration/Exploitation Trade-off Between Mind-Wandering and Goal-Directed Thinking*. The Oxford Handbook of Spontaneous Thought. https://doi.org/10.1093/oxfordhb/9780190464745.013.28

Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition*, *27*, 14–26. https://doi.org/10.1016/j.concog.2014.04.001

Uzzaman, S., & Joordens, S. (2011). The eyes know what you are thinking: eye movements as an objective measure of mind wandering. *Consciousness and Cognition*, *20*(4), 1882–1886. https://doi.org/10.1016/j.concog.2011.09.010